

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 14:09:36 ; Search time 22 Seconds
(without alignments)
1551.260 Million cell updates/sec

Title: US-09-492-028-2
Perfect score: 1828
Sequence: 1 MAGCCCLSAEEKESQRISAE.....VFAAVKDTILQLNLRBFNLV 355

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1828	100.0	355	2	A41534	GTP-binding protei
2	1780	97.4	355	2	A40891	GTP-binding protei
3	1538	84.1	359	1	RGMSQ	GTP-binding regula
4	1527	83.5	359	2	S30359	GTP-binding regula
5	1519	83.1	359	2	S71963	GTP-binding protei
6	1518	83.0	353	2	B40891	GTP-binding protei
7	1518	83.0	359	1	RGHUGY	GTP-binding regula
8	1517	83.0	359	2	S45700	G-alpha-11 protein
9	1512	82.7	359	1	RGMS11	GTP-binding regula
10	1502	82.2	359	2	S45699	GTP-binding regula
11	1408	77.0	355	2	T15288	hypothetical prote
12	1406	76.9	353	2	S34347	GTP-binding regula
13	1382.5	75.6	354	2	S33309	GTP-binding regula
14	1332.5	72.9	360	2	JN0115	GTP-binding regula
15	1022	55.9	374	2	B41534	GTP-binding protei
16	1013	55.4	374	2	A41096	GTP-binding regula
17	910.5	49.8	354	1	RGHUO2	GTP-binding regula
18	905.5	49.5	354	1	RGHYO2	GTP-binding regula
19	899.5	49.2	354	1	RGMSO2	GTP-binding regula
20	899.5	49.2	354	2	S27014	GTP-binding regula
21	897.5	49.1	354	1	RGFFO2	GTP-binding regula
22	895	49.0	355	2	I50238	Gi2 protein alpha-
23	894.5	48.9	354	1	RGTO2	GTP-binding regula
24	894.5	48.9	354	1	RGXLOA	GTP-binding regula
25	894	48.9	355	1	RGRTI2	GTP-binding regula
26	893	48.9	355	2	A61031	GTP-binding regula
27	893	48.9	355	2	S28158	GTP-binding regula
28	892.5	48.8	354	1	RGHUI1	GTP-binding regula
29	892.5	48.8	354	1	RGBOI1	GTP-binding regula

30	892.5	48.8	354	2	S24362	GTP-binding regula
31	892.5	48.8	354	2	S40508	GTP-binding regula
32	891.5	48.8	354	1	RGRTI1	GTP-binding regula
33	891.5	48.8	354	1	RGFFO1	GTP-binding regula
34	891	48.7	355	1	RGHUI2	GTP-binding regula
35	890.5	48.7	354	2	I50237	GTP-binding regula
36	888	48.6	355	1	RGMSI2	GTP-binding regula
37	884.5	48.4	354	1	RGXLI1	GTP-binding regula
38	884.5	48.4	354	2	S40509	G-protein - chicke
39	884.5	48.4	354	2	S28157	GTP-binding regula
40	883.5	48.3	354	2	A61035	GTP-binding regula
41	882	48.2	353	2	S71965	GTP-binding regula
42	880.5	48.2	354	1	RGHUO1	GTP-binding regula
43	877.5	48.0	354	1	RGHUI3	GTP-binding regula
44	877.5	48.0	354	1	RGRTI3	GTP-binding regula
45	873.5	47.8	354	2	S28159	GTP-binding regula

ALIGNMENTS

RESULT 1
A41534
GTP-binding protein alpha-14 chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 02-Feb-2001
C;Accession: A41534; E33833
R;Wilkie, T.M.; Scherle, P.A.; Strathmann, M.P.; Slepak, V.Z.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 88, 10049-10053, 1991
A;Title: Characterization of G-protein alpha subunits in the G-q class: expression in
A;Reference number: A41534; MUID:92052208; PMID:1946421
A;Accession: A41534
A;Molecule type: mRNA
A;Residues: 1-355 <WIL>
A;Cross-references: GB:M80631; NID:g193568; PIDN:AAA83222.1; PID:g193569
R;Strathmann, M.; Wilkie, T.M.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 86, 7407-7409, 1989
A;Title: Diversity of the G-protein family: sequences from five additional alpha-sub
A;Reference number: A33833; MUID:90017488; PMID:2508088
A;Accession: E33833
A;Molecule type: mRNA
A;Residues: 217-267 <STR>
A;Cross-references: GB:M57616; NID:g193380; PIDN:AAA63304.1; PID:g193381; GB:M26739
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; nucleotide binding; P-loop
F;42-49/Region: nucleotide-binding motif A (P-loop)
F;152-154/Region: GTP-binding SAK/L motif
F;270-273/Region: GTP-binding NKXD motif

Query Match 100.0%; Score 1828; DB 2; Length 355;
Best Local Similarity 100.0%; Pred. No. 1.7e-132;
Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGCCCLSAEEKESQRISAEIERHVRDKKDARRELKLLLTGTGESGKSTFIKQMRIHG	60
Db	1	MAGCCCLSAEEKESQRISAEIERHVRDKKDARRELKLLLTGTGESGKSTFIKQMRIHG	60
QY	61	SGYSDEDRKGF TKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTAL	120
Db	61	SGYSDEDRKGF TKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTAL	120
QY	121	SRDQVAAIKQLWLDPGIQCYDRRREYQLSDSAKYLYLTDIERIAMPFSVPTQQDVLVRV	180
Db	121	SRDQVAAIKQLWLDPGIQCYDRRREYQLSDSAKYLYLTDIERIAMPFSVPTQQDVLVRV	180
QY	181	PTGTIIEYPPDLENIIFRMVDVGGQSRERRKWIHCFSVTSIIFLVALUEYDQVLAECDN	240
Db	181	PTGTIIEYPPDLENIIFRMVDVGGQSRERRKWIHCFSVTSIIFLVALUEYDQVLAECDN	240
QY	241	ENRMEESKALFRTIITYPWFNLSSVILFLNKKDLEEKIMYSHLISYFFPEYTGPKQDVKA	300
Db	241	ENRMEESKALFRTIITYPWFNLSSVILFLNKKDLEEKIMYSHLISYFFPEYTGPKQDVKA	300

QY 301 ARDFILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILOLNLREFNLV 355
|||||
Db 301 ARDFILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILOLNLREFNLV 355

RESULT 2
A40891

GTP-binding protein GL1 alpha chain - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 02-Feb-2001
C;Accession: A40891

R;Nakamura, F.; Ogata, K.; Shiozaki, K.; Kameyama, K.; Ohara, K.; Haga, T.; Nukada, T.
J. Biol. Chem. 266, 12676-12681, 1991

A;Title: Identification of two novel GTP-binding protein alpha-subunits that lack apparent
A;Reference number: A40891; MUID:91286303; PMID:1905731

A;Accession: A40891

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-355 <NAK>

A;Cross-references: GB:D90335; NID:g217565; PIDN:BAAL4349.1; PID:g217566

C;Superfamily: GTP-binding regulatory protein Gs alpha chain

C;Keywords: GTP binding; nucleotide binding; P-loop

F;42-49/Region: nucleotide-binding motif A (P-loop)

F;152-154/Region: GTP-binding SAK/L motif

F;270-273/Region: GTP-binding NKXD motif

Query Match 97.4%; Score 1780; DB 2; Length 355;
Best Local Similarity 96.6%; Pred. No. 8e-129;
Matches 343; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAGCCLSAEKESQRIISAEIERHVRDKDARRELKLLLGTSGESGKSTFIKQMRIIHG 60
|||||
Db 1 MAGCCLSAEKESQRIISAEIERQLRRDKDARRELKLLLGTSGESGKSTFIKQMRIIHG 60

QY 61 SGYSDERKGF TKLVYQNIFTAMQAMIRAMDTLRIQYMCENKENAQIIREVEVDKVTAL 120
|||||
Db 61 SGYSDERKGF TKLVYQNIFTAMQAMIRAMDTLKIQYVCEQNKENAQLIREVEVDKVTAL 120

QY 121 SRDQVAAIKQLWDPGIECYDRRREYQLSDSAKYIYLTDIERIA MPSFVPTQQDVLVRV 180
|||||
Db 121 SRDQVEAIKQLWQDPGIECYDRRREYQLSDSAKYIYLTDIERIA MPFAFVPTQQDVLVRV 180

QY 181 PTTGIEYFPDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLValseyDQVLAECDN 240
|||||
Db 181 PTTGIEYFPDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLValseyDQVLAECDN 240

QY 241 ENRMEEKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFFPEYTGPKQDVKA 300
|||||
Db 241 ENRMEEKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFFPEYTGPKQDVKA 300

QY 301 ARDFILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILOLNLREFNLV 355
|||||
Db 301 ARDFILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILOLNLREFNLV 355

RESULT 3
RGMSQ

GTP-binding regulatory protein Gq alpha chain - mouse
N;Alternate names: guanine nucleotide binding protein Gq alpha chain; heterotrimeric G-protein
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Accession: A38414

R;Strathmann, M.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 87, 9113-9117, 1990

A;Title: G protein diversity: a distinct class of alpha subunits is present in vertebrates
A;Reference number: A38414; MUID:91067657; PMID:2123549

A;Accession: A38414

A;Molecule type: mRNA

A;Residues: 1-359 <STR>

A;Cross-references: GB:M55412; NID:g193501; PIDN:AAA63306.1; PID:g193502

C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay signals. The beta and gamma chains, required for GTPase activity, appear to be common to all G proteins. It is specific for each type of G protein.

C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction
F;46-53/Region: nucleotide-binding motif A (P-loop)
F;274-277/Region: GTP-binding NKXD motif
F;52/Binding site: GTP (Lys) #status predicted
F;183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 84.1%; Score 1538; DB 1; Length 359;
Best Local Similarity 82.3%; Pred. No. 2.8e-110;
Matches 289; Conservative 34; Mismatches 28; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRIISAEIERHVRDKDARRELKLLLGTSGESGKSTFIKQMRIIHGSGYS 64
|||||
Db 9 CCLSEAEKARRINDEIERHVRDKDARRELKLLLGTSGESGKSTFIKQMRIIHGSGYS 68

QY 65 DEDRKGF TKLVYQNIFTAMQAMIRAMDTLRIQYMCENKENAQIIREVEVDKVTALSRDQ 124
|||||
Db 69 DEDRKGF TKLVYQNIFTAMQAMIRAMDTLKIPIKYEHNKAHAQLREV DVEKVS AFENPY 128

QY 125 VAAIKQLWDPGIECYDRRREYQLSDSAKYIYLTDIERIA MPSFVPTQQDVLVRVPTTG 184
|||||
Db 129 VDAIKSLWDPGIECYDRRREYQLSDTKYYLNDLDRVADPSVLTQQDVLVRVPTTG 188

QY 185 IIEYFPDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLValseyDQVLAECDNENRM 244
|||||
Db 189 IIEYFPDLQSVIFRMVDVGGQSRERRKWIHCFENVTSIMFLValseyDQVLVESDNENRM 248

QY 245 EESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFFPEYTGPKQDVKAARDF 304
|||||
Db 249 EESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPQRDAQAAREF 308

QY 305 ILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILOLNLREFNLV 355
|||||
Db 309 ILKMFVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILOLNLKEYNLV 359

RESULT 4
S30359

GTP-binding regulatory protein G alpha chain, phospholipase C-activating - turkey
N;Alternate names: phospholipase C-activating G protein
C;Species: Meleagris gallopavo (common turkey)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 02-Feb-2001
C;Accession: S30359; S30360

R;Maurice, D.H.; Waldo, G.L.; Morris, A.J.; Nicholas, R.A.; Harden, T.K.
Biochem. J. 290, 765-770, 1993

A;Title: Identification of Galpha(11) as the phospholipase C-activating G-protein of
A;Reference number: S30359; MUID:93207527; PMID:8457205

A;Accession: S30359

A;Molecule type: mRNA

A;Residues: 1-359 <MAU>

A;Cross-references: GB:X73072; NID:g312254; PIDN:CAA51530.1; PID:g312255

A;Experimental source: blood

A;Accession: S30360

A;Molecule type: protein

A;Residues: 78-92;121-132;158-180;253-256;307-312;339-345;355-359 <MAW>

A;Experimental source: erythrocytes

C;Superfamily: GTP-binding regulatory protein Gs alpha chain

C;Keywords: GTP binding; nucleotide binding; P-loop

F;46-53/Region: nucleotide-binding motif A (P-loop)

F;156-158/Region: GTP-binding SAK/L motif

F;274-277/Region: GTP-binding NKXD motif

F;52/Binding site: GTP (Lys) #status predicted

F;183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 83.5%; Score 1527; DB 2; Length 359;
Best Local Similarity 82.1%; Pred. NO. 1.9e-109;
Matches 288; Conservative 36; Mismatches 27; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRIISAEIERHVRDKDARRELKLLLGTSGESGKSTFIKQMRIIHGSGYS 64
|||||
Db 9 CCLSDVEKESKRINAEIEKQLRRDKDARRELKLLLGTSGESGKSTFIKQMRIIHGSGYS 68

QY 65 DEDRKGF TKLVYQNIFTAMQAMIRAMDTLRIQYMCENKENAQIIREVEVDKVTALSRDQ 124

Db 69 EEDRKGF¹TKLVYQN¹IFTAMQSM¹IRAMET¹LKILYKYE¹QNKANAV¹LIREVD¹VEKVM¹TFEQPY¹ 128
QY 125 VAAIKQLW¹DPG¹IQECY¹DRRREY¹QLSDSAKY¹YLT¹DIERIAM¹PSFVPT¹QQDVL¹RVVP¹TTG¹ 184
Db 129 VSAIKTLW¹NDP¹G¹IQECY¹DRRREY¹QLSDSAKY¹YLS¹DVDRIAT¹PGYLP¹TQQDVL¹RVVP¹TTG¹ 188
QY 185 IIEY¹PF¹DL¹ENI¹IFRMV¹DVGG¹Q¹RSERR¹KW¹IHC¹FESV¹TSI¹IFL¹VALSEY¹DQVLA¹ECDNEN¹RM 244
Db 189 IIEY¹PF¹DL¹ENI¹IFRMV¹DVGG¹Q¹RSERR¹KW¹IHC¹FENV¹TSIM¹FLVALSEY¹DQVL¹VESDNEN¹RM 248
QY 245 EESKAL¹FRTI¹ITY¹PF¹LNSSV¹ILFL¹NKKDL¹LEEK¹IMY¹SHLISY¹FP¹EYTG¹PKQD¹VKAARD¹F 304
Db 249 EESKAL¹FRTI¹ITY¹PF¹LNSSV¹ILFL¹NKKDL¹LEDK¹ILY¹SHLV¹DYF¹PEFD¹GPQRDAQAARE¹F 308
QY 305 ILKLY¹QDQ¹NP¹DK¹EKVI¹YSH¹FTCAT¹DTENIR¹FVFAAV¹KD¹TILQLNL¹REFNLV¹ 355
Db 309 ILKMFV¹DLNP¹DS¹DKII¹YSH¹FTCAT¹DTENIR¹FVFAAV¹KD¹TILQLNL¹KEYNLV¹ 359

RESULT 5
S71963
GTP-binding protein alpha-q - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 02-Feb-2001
C;Accession: S71963
R;Johnson, G.J.; Leis, L.A.; Dunlop, P.C.
Biochem. J. 318, 1023-1031, 1996
A;Title: Specificity of G-alpha(q) and G-alpha(11) gene expression in platelets and erythrocytes
A;Reference number: S71963; MUID:96433124; PMID:8836152
A;Accession: S71963
A;Molecule type: mRNA
A;Residues: 1-359 <JOH>
A;Cross-references: EMBL:L76256; NID:g1478071; PIDN:AAB39498.1; PID:g1478072
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; nucleotide binding; p-loop; signal transduction
F;46-53/Region: nucleotide-binding motif A (P-loop)
F;274-277/Region: GTP-binding NKXD motif

Query Match 83.1%; Score 1519; DB 2; Length 359;
Best Local Similarity 81.2%; Pred. No. 7.9e-109;
Matches 285; Conservative 36; Mismatches 30; Indels 0; Gaps 0;

QY 5 CCLSAE¹EKE¹SQRISA¹EIERH¹VR¹DK¹DARREL¹KLL¹LTG¹ESGK¹STFI¹KQM¹RIIHGSGYS¹ 64
Db 9 CCLSEE¹AKEARR¹INDEI¹ERQL¹RRDK¹RDARREL¹KLL¹LTG¹ESGK¹STFI¹KQM¹RIIHGSGYS¹ 68
QY 65 DEDRK¹GFT¹KL¹VYQ¹NI¹FTAMQ¹AMIRAM¹DTL¹RIQ¹YMC¹EQNK¹ENAI¹IREVE¹VDK¹VTALSRDQ¹ 124
Db 69 DEDRK¹GFT¹KL¹VYQ¹NI¹FTAMQ¹AMIRAM¹DTL¹KIPY¹KYEH¹NKAHAQL¹VREV¹DVEKVS¹AFENPY¹ 128
QY 125 VAAIKQLW¹DPG¹IQECY¹DRRREY¹QLSDSAKY¹YLT¹DIERIAM¹PSFVPT¹QQDVL¹RVVP¹TTG¹ 184
Db 129 VDAIKSLW¹NDP¹G¹IQECY¹DRRREY¹QLSDSTKY¹YLN¹DLDRVAD¹PAYLP¹TQQDVL¹RVVP¹TTG¹ 188
QY 185 IIEY¹PF¹DL¹ENI¹IFRMV¹DVGG¹Q¹RSERR¹KW¹IHC¹FESV¹TSI¹IFL¹VALSEY¹DQVLA¹ECDNEN¹RM 244
Db 189 IIEY¹PF¹DL¹QSVI¹FRMV¹DVGG¹Q¹RSERR¹KW¹IHC¹FENV¹TSIM¹FLVALSEY¹DQVL¹VESDNEN¹RM 248
QY 245 EESKAL¹FRTI¹ITY¹PF¹LNSSV¹ILFL¹NKKDL¹LEEK¹IMY¹SHLISY¹FP¹EYTG¹PKQD¹VKAARD¹F 304
Db 249 EESKAL¹FRTI¹ITY¹PF¹LNSSV¹ILFL¹NKKDL¹LEEK¹IMY¹SHLV¹DYF¹PEFD¹GPQRDAQAARE¹F 308
QY 305 ILKLY¹QDQ¹NP¹DK¹EKVI¹YSH¹FTCAT¹DTENIR¹FVFAAV¹KD¹TILQLNL¹REFNLV¹ 355
Db 309 ILKMFV¹DLNP¹DS¹DKII¹YSH¹FTCAT¹DTENIR¹FVFAAV¹KD¹TILQLNL¹KEYNAV¹ 359

RESULT 6
B40891
GTP-binding protein GL2 alpha chain - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 02-Feb-2001

C;Accession: B40891
R;Nakamura, F.; Ogata, K.; Shiozaki, K.; Kameyama, K.; Ohara, K.; Haga, T.; Nukada, T
J. Biol. Chem. 266, 12676-12681, 1991
A;Title: Identification of two novel GTP-binding protein alpha-subunits that lack app
A;Reference number: A40891; MUID:91286303; PMID:1905731
A;Accession: B40891
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-353 <NAK>
A;Cross-references: GB:D90336; NID:g217567; PIDN:BAA14350.1; PID:g217568
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; nucleotide binding; p-loop
F;40-47/Region: nucleotide-binding motif A (P-loop)
F;150-152/Region: GTP-binding SAK/L motif
F;268-271/Region: GTP-binding NKXD motif

Query Match 83.0%; Score 1518; DB 2; Length 353;
Best Local Similarity 81.5%; Pred. No. 9.2e-109;
Matches 286; Conservative 37; Mismatches 28; Indels 0; Gaps 0;

QY 5 CCLSAE¹EKE¹SQRISA¹EIERH¹VR¹DK¹DARREL¹KLL¹LTG¹ESGK¹STFI¹KQM¹RIIHGSGYS¹ 64
Db 3 CCLSD¹EVKESKRINAE¹IEQL¹RRDK¹RDARREL¹KLL¹LTG¹ESGK¹STFI¹KQM¹RIIHGAGYS¹ 62
QY 65 DEDRK¹GFT¹KL¹VYQ¹NI¹FTAMQ¹AMIRAM¹DTL¹RIQ¹YMC¹EQNK¹ENAI¹IREVE¹VDK¹VTALSRDQ¹ 124
Db 63 EEDK¹RGFT¹KL¹VYQ¹NI¹FTAMQ¹AMIRAM¹ETL¹KILY¹KYE¹QNKANAL¹LIREVD¹VEKVT¹FEHRY¹ 122
QY 125 VAAIKQLW¹DPG¹IQECY¹DRRREY¹QLSDSAKY¹YLT¹DIERIAM¹PSFVPT¹QQDVL¹RVVP¹TTG¹ 184
Db 123 VSAIKTLW¹NDP¹G¹IQECY¹DRRREY¹QLSDSAKY¹YLT¹DVDRAT¹SGYLP¹TQQDVL¹RVVP¹TTG¹ 182
QY 185 IIEY¹PF¹DL¹ENI¹IFRMV¹DVGG¹Q¹RSERR¹KW¹IHC¹FESV¹TSI¹IFL¹VALSEY¹DQVLA¹ECDNEN¹RM 244
Db 183 IIEY¹PF¹DL¹ENI¹IFRMV¹DVGG¹Q¹RSERR¹KW¹IHC¹FENV¹TSIM¹FLVALSEY¹DQVL¹VESDNEN¹RM 242
QY 245 EESKAL¹FRTI¹ITY¹PF¹LNSSV¹ILFL¹NKKDL¹LEEK¹IMY¹SHLISY¹FP¹EYTG¹PKQD¹VKAARD¹F 304
Db 243 EESKAL¹FRTI¹ITY¹PF¹LNSSV¹ILFL¹NKKDL¹LEDK¹ILH¹SHLV¹DYF¹PEFD¹GPQRDAQAARE¹F 302
QY 305 ILKLY¹QDQ¹NP¹DK¹EKVI¹YSH¹FTCAT¹DTENIR¹FVFAAV¹KD¹TILQLNL¹REFNLV¹ 355
Db 303 ILKMFV¹DLNP¹DS¹DKII¹YSH¹FTCAT¹DTENIR¹FVFAAV¹KD¹TILQLNL¹KEYNLV¹ 353

RESULT 7
RGHUGY
GTP-binding regulatory protein Gy alpha chain - human
N;Alternate names: guanine nucleotide binding protein Gy alpha chain; heterotrimeric
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Accession: A39394
R;Jiang, M.; Pandey, S.; Tran, V.T.; Fong, H.K.W.
Proc. Natl. Acad. Sci. U.S.A. 88, 3907-3911, 1991
A;Title: Guanine nucleotide-binding regulatory proteins in retinal pigment epithelial
A;Reference number: A39394; MUID:91219481; PMID:1902575
A;Accession: A39394
A;Molecule type: mRNA
A;Residues: 1-359 <JIA>
A;Cross-references: GB:M69013; NID:g183690; PIDN:AAA58624.1; PID:g183691
C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that re
ains. The beta and gamma chains, required for GTPase activity, appear to be common to
rase; it is specific for each type of G protein.
C;Genetics:

A;Gene: GDB:GNA11
A;Cross-references: GDB:132587; OMIM:139313
A;Map position: 19p13.3-19p13.3
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; heterotrimer; nucleotide binding; p-loop; signal transductio
F;46-53/Region: nucleotide-binding motif A (P-loop)
F;274-277/Region: GTP-binding NKXD motif
F;52/Binding site: GTP (Lys) #status predicted
F;183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Biochem. J. 292, 333-341, 1993
A;Title: The molecular cloning of the squid (Loligo forbesi) visual Gq-alpha subunit and
A;Reference number: S33309; MUID:93277493; PMID:8503868
A;Accession: S33309
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-354 <RYB>
A;Cross-references: EMBL:L10289
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; nucleotide binding; p-loop
F;40-47/Region: nucleotide-binding motif A (p-loop)
F;150-152/Region: GTP-binding SAK/L motif
F;269-272/Region: GTP-binding NKXD motif

Query Match 75.6%; Score 1382.5; DB 2; Length 354;
Best Local Similarity 74.4%; Pred. No. 2.2e-98;
Matches 262; Conservative 37; Mismatches 52; Indels 1; Gaps 1;

QY 5 CCLSAEKEQSQRISAEIERHVRDKDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 64
Db 3 CCLSEEAKEQKRINQIEKQLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 62
QY 65 DEDRKGF TKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ 124
Db 63 EEDRKGF EKIVYQNIFSAIQTLIAAMETLSLEYKDPNSNHEAFLNSIDADSADIFEDGH 122
QY 125 VAAIKQLWLDPGIQECYDRRREYQLSDSAKYLLTDIERIAMP SFVPTQQDVLVRVPTTG 184
Db 123 VTAIKGCWTDPGMQECYDRRREYQLTDSAKYYLDDVERIHEPGYIPTLQDILVRVPTTG 182
QY 185 IIEYPPDLNIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDN-ENR 243
Db 183 IIEYPPDLYSIIFRMVDVGGQSRERRKWIHCFENVTSINFLVALSEYDQVLVESDNEENR 242
QY 244 MEESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFP EYTGPKQDVKAARD 303
Db 243 MEESKALFRTIITYPWFQNSVILFLNKKDLLEEKIMTSHLADYFPDYDGPKCDYEAARE 302
QY 304 FILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQNLNREFNLV 355
Db 303 FMMSYMDLNEDKEMLYHYHTCATDTENIRFVFAAVKDTILQNLNKEYNLV 354

RESULT 14
JN0115
GTP-binding regulatory protein dgq alpha chain - fruit fly (Drosophila melanogaster)
N;Alternate names: dgq protein
C;Species: Drosophila melanogaster
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 02-Feb-2001
C;Accession: JN0115
R;Lee, Y.J.; Dobbs, M.B.; Verardi, M.L.; Hyde, D.R.
Neuron 5, 889-898, 1990
A;Title: dgq: a drosophila gene encoding a visual system-specific G alpha molecule.
A;Reference number: JN0115; MUID:91097801; PMID:2125225
A;Accession: JN0115
A;Molecule type: DNA
A;Residues: 1-360 <LEE>
C;Genetics:
A;Gene: dgq
A;Cross-references: FlyBase:FBgn0004435
A;Introns: 40/1; 101/3; 153/2; 196/2; 239/3; 298/1; 334/1
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: alternative splicing; GTP binding; nucleotide binding; p-loop
F;1-360/Product: GTP-binding regulatory protein dgq alpha chain I #status predicted <MAT
F;1-297,333-360/Product: GTP-binding regulatory protein dgq alpha chain II #status predi
F;40-47/Region: nucleotide-binding motif A (p-loop)
F;150-152/Region: GTP-binding SAK/L motif
F;268-271/Region: GTP-binding NKXD motif

Query Match 72.9%; Score 1332.5; DB 2; Length 360;
Best Local Similarity 71.4%; Pred. No. 1.5e-94;
Matches 255; Conservative 39; Mismatches 56; Indels 7; Gaps 1;

QY 5 CCLSAEKEQSQRISAEIERHVRDKDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 64
Db 3 CCLSEEAKEQKRINQIEKQLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 62
QY 65 DEDRKGF TKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ 124
Db 63 DEDKRGYIKLVFQNFMAQSMIKAMDMLKISYGQGHSELADLVMSIDYETVTFEDPY 122
QY 125 VAAIKQLWLDPGIQECYDRRREYQLSDSAKYLLTDIERIAMP SFVPTQQDVLVRVPTTG 184
Db 123 LNAIKTLWDDAGIQECYDRRREYQLTDSAKYYLSDLARIEQAADYLPTEQDILRARVPTTG 182
QY 185 IIEYPPDLNIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDNENRM 244
Db 183 ILEYPPDLDGIVFRMVDVGGQSRERRKWIHCFENVTSIIFLVALSEYDQILFESDNEENRM 242
QY 245 EESKALFRTIITYPWFNLSSVILFLNKK-----DLLEEKIMYSHLISYFP EYTGPKQD 297
Db 243 EESKALFRTIITYPWFQNSVILFLNKKDLLEEKIDLLEEKIMYSHLVDYFFPEYDGPQD 302
QY 298 VKAARDFILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQNLNREFNL 354
Db 303 HAAAKQFVLKKYLACNPDPERQCYSHFTTATDTENIKLVCAVKDTIMQNALKEFNL 359

Query Match 55.9%; Score 1022; DB 2; Length 374;
Best Local Similarity 54.8%; Pred. No. 9.6e-71;
Matches 201; Conservative 60; Mismatches 92; Indels 14; Gaps 3;

QY 3 GCC--CLSAEKEQSQRISAEIERHVRDKDARRELKLLLLGTGESGKSTFIKQMRIIHG 60
Db 8 GCCPWLCTEEKTAARIDQEI NRILLEQKKQEREELKLLLLGPGESGKSTFIKQMRIIHG 67
QY 61 SCYSDEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTAL 120
Db 68 VGyseEDRRAFRLLIYQNIFVSMQAMIDAMDRLOIPFSRPDSKQHASLVMTQDPYKVSTF 127
QY 121 SRDQVAAIKQLWLDPGIQECYDRRREYQLSDSAKYLLTDIERIAMP SFVPTQQDVLVRV 180
Db 128 EKPYAVAMQYLRDAGIRACYERRRREFHLLDSAVYLSHLERISEDSYIPTAQDVLRSRM 187
QY 181 PTTGIIEYPFDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDN 240
Db 188 PTTGNEYCFSVKKTCLRIVDVGQSRERRKWIHCFENVIALIYLASLSEYDQCLEENDQ 247
QY 241 ENRMEESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFP EYTGPKQDVKA 300
Db 248 ENRMEESLALFSTILELPWFKSTSVILFLNKTILEDKIHTSHLATYFP SQGRRDAEA 307
QY 301 ARDFILKLY-----QD--QNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQNL 348
Db 308 AKSFILDMYARVYASCAEPQDGGKGRARRFFAHTTCATDTQSVRSVFKDVRDSVLARY 367

Qy 349 LREFNLV 355
| | | |
Db 368 LDEINLL 374

Search completed: February 3, 2003, 14:12:10
Job time : 23 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 14:03:46 ; Search time 38 Seconds
(without alignments)
1244.841 Million cell updates/sec

Title: US-09-492-028-2
Perfect score: 1828
Sequence: 1 MAGCCCLSAEKEQSRIAE.....VFAAVKDTILQLNLFNLY 355

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	A_Geneseq_101002:*
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1828	100.0	355	21 AAB15026	Mouse TC-Galpha14
2	1771	96.9	355	22 AAB99076	Human G-protein al
3	1771	96.9	355	23 ABB09285	G protein-coupled
4	1537	84.1	353	23 ABG68591	Mouse G protein al
5	1537	84.1	359	23 ABG68592	Mouse G protein al
6	1535	84.0	359	23 AAM48816	Murine G-protein 6
7	1530	83.7	353	23 ABG68588	Mouse G protein al
8	1530	83.7	359	23 ABG68594	Mouse G protein al
9	1527	83.5	353	23 ABG68585	Mouse G protein al
10	1527	83.5	353	23 ABG68600	Human G protein al

11	1527	83.5	359	23 ABG68584	Mouse G protein al
12	1527	83.5	359	23 ABG68604	Human G protein al
13	1526	83.5	353	23 ABG68587	Mouse G protein al
14	1526	83.5	359	23 ABG68586	Mouse G protein al
15	1524	83.4	359	23 ABG68610	Human G protein al
16	1524	83.4	359	23 ABB09282	G protein-coupled
17	1522	83.3	359	23 ABG68608	Human G protein al
18	1520	83.2	353	23 AAM48817	Murine G-protein 6
19	1520	83.2	353	23 AAM48818	Murine G-protein 6
20	1520	83.2	353	23 AAM48819	Murine G-protein 6
21	1519	83.1	353	22 AAB99072	Human G-protein al
22	1519	83.1	353	23 ABG68599	Human G-protein al
23	1519	83.1	353	23 ABB09281	G protein-coupled
24	1519	83.1	359	23 ABG68593	Mouse G protein al
25	1519	83.1	359	23 ABG68595	Mouse G protein al
26	1519	83.1	359	23 ABG68598	Human G protein al
27	1518	83.0	353	23 ABG68597	Mouse G protein al
28	1518	83.0	359	22 AAB99071	Human G-protein al
29	1518	83.0	359	22 AAB99073	Human G-protein al
30	1517	83.0	359	20 AAY52705	Human G-alpha-11 p
31	1517	83.0	359	20 AAY29789	Human G-alpha-11 p
32	1516	82.9	359	23 ABG68603	Human G protein al
33	1516	82.9	359	23 ABG68605	Human G protein al
34	1515	82.9	353	23 ABG68607	Human G protein al
35	1515	82.9	359	23 ABG68609	Human G protein al
36	1512	82.7	353	23 ABG68589	Mouse G protein al
37	1512	82.7	359	23 ABG68596	Mouse G protein al
38	1509	82.5	353	23 ABG68601	Human G protein al
39	1509	82.5	359	23 ABG68606	Human G protein al
40	1506	82.4	1276	20 AAY49127	phCaR/hmGluR2*Gq15
41	1506	82.4	1276	23 AAO15093	Human phCaR/hmGluR
42	1506	82.4	1394	20 AAY49129	pmGluR2/CaR*Galpha
43	1506	82.4	1394	23 AAO15095	Human pmGluR2-CaR*
44	1506	82.4	1397	20 AAY49134	pmGluR2/CaR*Galpha
45	1506	82.4	1397	23 AAO15100	Human pmGluR2-CaR*

ALIGNMENTS

RESULT 1	
AAB15026	
ID	AAB15026 standard; Protein; 355 AA.
XX	
AC	AAB15026;
XX	
DT	18-DEC-2000 (first entry)
XX	
DE	Mouse TC-Galpha14 protein.
XX	
KW	Mouse; TC-Galpha14; sensory signal; taste.
XX	
OS	Mus sp.
XX	
PN	WO200044929-A2.
XX	
PD	03-AUG-2000.
XX	
PF	26-JAN-2000; 2000WO-US02217.
XX	
PR	27-JAN-1999; 99US-0117367.
XX	
PA	(REGC) UNIV CALIFORNIA.
XX	
PI	Zuker CS;
XX	
DR	WPI; 2000-499336/44.
XX	
XX	N-PSDB; AAA73799.

Assaying for compounds that modulate sensory signalling in taste cells, by determining interactions between the compounds and a sensory cell specific G-protein alpha subunit polypeptide -

PS Claim 1; Page 61-62; 67pp; English.

XX

CC The present invention relates to a method for identifying a compound

CC that modulates sensory signalling in taste cells, comprising contacting

CC the compound with a sensory cell specific G-protein alpha subunit

CC polypeptide. Such a polypeptide is the present sequence, mouse

CC TC-Galpal4 protein. The compounds identified by the present method

CC may be used by the food and pharmaceutical industries to customize

CC taste as additives for food or medicines so that they taste different

CC when eaten.

XX

SQ Sequence 355 AA;

Query Match 100.0%; Score 1828; DB 21; Length 355;

Best Local Similarity 100.0%; Pred. No. 6.9e-173;

Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGCCCLSAEEKESQRISAEIERHVRRDKKDARRELKLLLLGTGESGKSTFIKQMRIIHG 60

Db 1 MAGCCCLSAEEKESQRISAEIERHVRRDKKDARRELKLLLLGTGESGKSTFIKQMRIIHG 60

QY 61 SGYSDEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTAL 120

Db 61 SGYSDEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTAL 120

QY 121 SRDQVAAIKQLWLDPGIQECYDRRREYQLSDSAKYYLTDIERIAMPVSFVPTQQDVLVRV 180

Db 121 SRDQVAAIKQLWLDPGIQECYDRRREYQLSDSAKYYLTDIERIAMPVSFVPTQQDVLVRV 180

QY 181 PTTGIIEYFPDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIIFLVALSEYDQVLAECDN 240

Db 181 PTTGIIEYFPDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIIFLVALSEYDQVLAECDN 240

QY 241 ENRMEESKALFRITITYPWFNLSSVILFLNKKDLLLEEKIMYSHLISYFPETGPKQDVKA 300

Db 241 ENRMEESKALFRITITYPWFNLSSVILFLNKKDLLLEEKIMYSHLISYFPETGPKQDVKA 300

QY 301 ARDFILKLYQDQNPDKVKVIYSHFTCATDTENIRFVFAAVKDTILQLNLRREFNLV 355

Db 301 ARDFILKLYQDQNPDKVKVIYSHFTCATDTENIRFVFAAVKDTILQLNLRREFNLV 355

RESULT 2

AAB99076

ID AAB99076 standard; Protein; 355 AA.

XX

AC AAB99076;

XX

DT 23-AUG-2001 (first entry)

XX

DE Human G-protein alpha subunit 14.

XX

KW G-protein coupled receptor; GPCR; GnRH receptor; disease treatment;

KW gonadotrophin releasing; hormone receptor; hormone dependent cancer;

KW human; catfish; goldfish; cow; sheep; horse; fruitfly; pig; rat; mouse;

KW gene therapy.

XX

OS Homo sapiens.

XX

PN WO200136446-A2.

XX

PD 25-MAY-2001.

XX

PF 17-NOV-2000; 2000WO-GB04385.

XX

PR 17-NOV-1999; 99GB-0027215.

XX

PA (UYBR-) UNIV BRISTOL.

XX

PI McArdle CA;

XX

DR WPI; 2001-355607/37.

XX

PT Use of a vector encoding G-protein coupled receptors for manufacturing

PT medicaments for treating cancer, diseases of cardiovascular system,

PT nervous system, digestive system, immune system, or muscle diseases

XX

PS Disclosure; Fig 19; 78pp; English.

XX

CC The present invention describes a prodrug comprising a vector encoding a

CC G-protein coupled receptor (GPCR). This can be used in the treatment of

CC diseases, including hormone-dependent cancers, cardiovascular, nervous

CC system, digestive system, immune system, respiratory, skeletal,

CC endocrine, sensory and muscle diseases and disorders. The present

CC sequence is a protein described in the exemplification of the invention.

XX

SQ Sequence 355 AA;

Query Match 96.9%; Score 1771; DB 22; Length 355;

Best Local Similarity 96.3%; Pred. No. 3.2e-167;

Matches 342; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAGCCCLSAEEKESQRISAEIERHVRRDKKDARRELKLLLLGTGESGKSTFIKQMRIIHG 60

Db 1 MAGCCCLSAEEKESQRISAEIERQLRRDKKDARRELKLLLLGTGESGKSTFIKQMRIIHG 60

QY 61 SGYSDEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTAL 120

Db 61 SGYSDEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYVCEQNKENAQIIREVEVDKVSML 120

QY 121 SRDQVAAIKQLWLDPGIQECYDRRREYQLSDSAKYYLTDIERIAMPVSFVPTQQDVLVRV 180

Db 121 SREQVEAIKQLWQDPGQECYDRRREYQLSDSAKYYLTDIDRIATPVSFVPTQQDVLVRV 180

QY 181 PTTGIIEYFPDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIIFLVALSEYDQVLAECDN 240

Db 181 PTTGIIEYFPDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIIFLVALSEYDQVLAECDN 240

QY 241 ENRMEESKALFRITITYPWFNLSSVILFLNKKDLLLEEKIMYSHLISYFPETGPKQDVKA 300

Db 241 ENRMEESKALFRITITYPWFNLSSVILFLNKKDLLLEEKIMYSHLISYFPETGPKQDVRA 300

QY 301 ARDFILKLYQDQNPDKVKVIYSHFTCATDTENIRFVFAAVKDTILQLNLRREFNLV 355

Db 301 ARDFILKLYQDQNPDKVKVIYSHFTCATDTENIRFVFAAVKDTILQLNLRREFNLV 355

RESULT 3

ABB09285

ID ABB09285 standard; Protein; 355 AA.

XX

AC ABB09285;

XX

DT 10-JUL-2002 (first entry)

XX

DE G protein-coupled receptor (GPCR) >G(Y-14) SEQ ID NO:31.

XX

KW Target activated nucleic acid biosensor; signalling moiety; GPCR;

KW nucleic acid sensor; detection; engineering; drug optimisation;

KW G protein-coupled receptor.

XX

OS Homo sapiens.

XX

PN WO200222882-A2.

XX

PD 21-MAR-2002.

XX

PF 13-SEP-2001; 2001WO-US28835.

XX

PR 13-SEP-2000; 2000US-232454P.

XX

PA (ARCH-) ARCHEMIX CORP.

XX

PI Stanton M, Epstein D, Hamaguchi N;

XX

DR WPI; 2002-393977/42.

XX	Mouse G protein alpha sub-unit q family variant #8.
DE	
XX	
KW	G_q protein; sensory signaling; chemoreceptor; tastant; olfactant;
KW	pheromone; G protein alpha sub-unit; q family; G alpha q.
XX	
XX	Mus musculus.
OS	
XX	
PN	WO200236622-A2.
XX	
PD	10-MAY-2002.
XX	
PF	24-OCT-2001; 2001WO-US32619.
XX	
PR	30-OCT-2000; 2000US-243770P.
XX	
PA	(SENO-) SENOMYX INC.
XX	
PI	Yao Y, Xu H;
XX	
DR	WPI; 2002-519234/55.
XX	
PT	New G-alpha-q protein variants, useful for analyzing and discovering
PT	agonists or antagonists of chemoreceptors, such as G protein coupled
PT	receptors involved in sensing of tastants, olfactants or pheromones -
XX	
PS	Claim 13; Page 26; 32pp; English.
XX	
CC	The invention describes an isolated variant of a G_q protein, which
CC	exhibits increased promiscuity relative to the corresponding G_q protein.
CC	The variant is used to identify a compound that modulates sensory
CC	signaling in sensory cells and to identify a compound that interacts
CC	with the G_q variant protein. The G_q protein variant is useful for
CC	analysing and discovering agonists or antagonists of chemoreceptors, such
CC	as G protein coupled receptors involved in sensing of tastants,
CC	olfactants or pheromones. This is the amino acid sequence of a G protein
CC	alpha sub-unit q family (G alpha q) variant that can functionally couple
CC	to sensory cell receptors such as taste GPCR's (G protein-coupled
CC	receptors) and olfactory GPCR's in an overly promiscuous manner.
XX	
SQ	Sequence 359 AA;

```

ID AAM48816 standard; Protein; 359 AA.
XX
XX AC AAM48816;
XX
XX DT 24-APR-2002 (first entry)
XX
XX DE Murine G-protein 6qi4myr.
XX
XX KW Mouse; G-protein coupled receptor modulator; G-protein 6qi4myr;
XX KW signal transduction.
XX
XX OS Mus musculus.
XX
XX PN WO200204665-A2.
XX
XX PD 17-JAN-2002.
XX
XX PF 05-JUL-2001; 2001WO-EP07667.
XX
XX PR 08-JUL-2000; 2000DE-1033353.
XX
XX PA (AVET ) AVENTIS PHARMA DEUT GMBH.
XX
XX PI Kostenis E;
XX
XX DR WPI; 2002-148182/19.
XX DR N-PSDB; ABA97518.
XX
XX PT Identifying compounds that modify activity of signal transduction
XX PT pathways, useful potentially as therapeutic agents, by screening with
XX PT cells that contain hybrid G proteins -
XX
XX PS Claim 22; Page 24-25; 34pp; German.
XX
XX CC The present invention relates to a method of identifying compounds which
XX CC are capable of modifying the activity of a signal transduction pathway
XX CC which is dependent upon a G-protein coupled receptor. These compounds may
XX CC include the novel G-proteins 6qi4myr (shown here), 6qi5myr, 6qi4, 6qs5
XX CC and Galphal6.
XX
XX SQ Sequence 359 AA;

```


||||| 249 EESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDFYFPEYDGPQORDAQAAAREF 308

QY 305 ILKLYQDQNPDKKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355

Db 309 ILKMFVDLNPDSDKIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 9

ABG68585

ID ABG68585 standard; Protein; 353 AA.

XX

AC ABG68585;

XX

DT 07-OCT-2002 (first entry)

DE Mouse G protein alpha sub-unit q family variant #1.

XX

KW G_q protein; sensory signaling; chemoreceptor; tastant; olfactant;

KW pheromone; G protein alpha sub-unit; q family; G alpha q.

XX

OS Mus musculus.

XX

XX WO200236622-A2.

PN

XX 10-MAY-2002.

PD

XX

PF 24-OCT-2001; 2001WO-US32619.

XX

XX 30-OCT-2000; 2000US-243770P.

PR

XX

PA (SENO-) SENOMYX INC.

XX

XX

PI Yao Y, Xu H;

XX

DR WPI; 2002-519234/55.

XX

XX New G-alpha-q protein variants, useful for analyzing and discovering agonists or antagonists of chemoreceptors, such as G protein coupled receptors involved in sensing of tastants, olfactants or pheromones

PS

XX Claim 13; Page 23; 32pp; English.

XX

CC The invention describes an isolated variant of a G_q protein, which exhibits increased promiscuity relative to the corresponding G_q protein.

CC

CC The variant is used to identify a compound that modulates sensory signaling in sensory cells and to identify a compound that interacts with the G_q variant protein. The G_q protein variant is useful for analysing and discovering agonists or antagonists of chemoreceptors, such as G protein coupled receptors involved in sensing of tastants, olfactants or pheromones. This is the amino acid sequence of a G protein alpha sub-unit q family (G alpha q) variant that can functionally couple to sensory cell receptors such as taste GPCR's (G protein-coupled receptors) and olfactory GPCR's in an overly promiscuous manner.

XX

SQ Sequence 353 AA;

Query Match 83.5%; Score 1527; DB 23; Length 353;

Best Local Similarity 81.8%; Pred. No. 5.4e-143;

Matches 287; Conservative 35; Mismatches 29; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRIAEIERHVRRDKDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 64

Db 3 CCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 62

QY 65 DEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAIIREVEVDKVTALSRDQ 124

Db 63 DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLREVVDVEKVSFENPY 122

QY 125 VAAIKQLWDPGIECYDRRREYQLSDSAKYIYLTDIERIA MPSFVPTQDVLVRVPPTG 184

Db 123 VDAIKSLWNDPGIECYDRRREYQLSDSTKYIYLNLDLRVADPSYLP TQDVLVRVPPTG 182

QY 185 IIEYPFDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDNENRM 244

Db 183 IIEYPFDLQSVIFRMVDVGGQSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 242

QY 245 EESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAARDF 304

Db 243 EESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDFYFPEYDGPQORDAQAAAREF 302

QY 305 ILKLYQDQNPDKKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355

Db 303 ILKMFVDLNPDSDKIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 353

RESULT 10

ABG68600

ID ABG68600 standard; Protein; 353 AA.

XX

AC ABG68600;

XX

DT 07-OCT-2002 (first entry)

DE Human G protein alpha sub-unit q family variant #2.

XX

KW G_q protein; sensory signaling; chemoreceptor; tastant; olfactant;

KW pheromone; G protein alpha sub-unit; q family; G alpha q.

XX

OS Homo sapiens.

XX

PN WO200236622-A2.

XX

PD 10-MAY-2002.

XX

XX 24-OCT-2001; 2001WO-US32619.

PF

XX 30-OCT-2000; 2000US-243770P.

PR

XX

PA (SENO-) SENOMYX INC.

XX

PI Yao Y, Xu H;

XX

DR WPI; 2002-519234/55.

XX

XX New G-alpha-q protein variants, useful for analyzing and discovering agonists or antagonists of chemoreceptors, such as G protein coupled receptors involved in sensing of tastants, olfactants or pheromones

PS

XX Claim 13; Page 29; 32pp; English.

XX

CC The invention describes an isolated variant of a G_q protein, which exhibits increased promiscuity relative to the corresponding G_q protein.

CC

CC The variant is used to identify a compound that modulates sensory signaling in sensory cells and to identify a compound that interacts with the G_q variant protein. The G_q protein variant is useful for analysing and discovering agonists or antagonists of chemoreceptors, such as G protein coupled receptors involved in sensing of tastants, olfactants or pheromones. This is the amino acid sequence of a G protein alpha sub-unit q family (G alpha q) variant that can functionally couple to sensory cell receptors such as taste GPCR's (G protein-coupled receptors) and olfactory GPCR's in an overly promiscuous manner.

XX

SQ Sequence 353 AA;

Query Match 83.5%; Score 1527; DB 23; Length 353;

Best Local Similarity 81.8%; Pred. No. 5.4e-143;

Matches 287; Conservative 35; Mismatches 29; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRIAEIERHVRRDKDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 64

Db 3 CCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSDYS 62

QY 65 DEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAIIREVEVDKVTALSRDQ 124

Db 63 DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLREVVDVEKVSADFVPD 122

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 14:10:47 ; Search time 16 Seconds
(without alignments)
652.821 Million cell updates/sec

Title: US-09-492-028-2
Perfect score: 1828
Sequence: 1 MAGCCCLSAEKEKESQRISAE.....VFAAVKDTILQLNLFENLV 355

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1042	57.0	374	4 US-09-442-349A-29	Sequence 29, Appl
2	1042	57.0	374	4 US-09-442-349A-31	Sequence 31, Appl
3	1041	56.9	374	4 US-09-442-349A-21	Sequence 21, Appl
4	1041	56.9	374	4 US-09-442-349A-23	Sequence 23, Appl
5	1041	56.9	374	4 US-09-442-349A-25	Sequence 25, Appl
6	1041	56.9	374	4 US-09-442-349A-30	Sequence 30, Appl
7	1041	56.9	374	4 US-09-442-349A-32	Sequence 32, Appl
8	1041	56.9	374	4 US-09-442-349A-34	Sequence 34, Appl
9	1040	56.9	374	4 US-09-442-349A-2	Sequence 2, Appl
10	1040	56.9	374	4 US-09-442-349A-22	Sequence 22, Appl
11	1040	56.9	374	4 US-09-442-349A-24	Sequence 24, Appl
12	1040	56.9	374	4 US-09-442-349A-33	Sequence 33, Appl
13	1037	56.7	374	4 US-09-442-349A-26	Sequence 26, Appl
14	1037	56.7	374	4 US-09-442-349A-28	Sequence 28, Appl
15	1036	56.7	374	4 US-09-442-349A-20	Sequence 20, Appl
16	1036	56.7	374	4 US-09-442-349A-27	Sequence 27, Appl
17	1022	55.9	374	3 US-08-878-801-4	Sequence 4, Appl
18	1021	55.9	374	4 US-09-442-349A-14	Sequence 14, Appl
19	1021	55.9	374	4 US-09-442-349A-16	Sequence 16, Appl
20	1020	55.8	365	4 US-09-442-349A-108	Sequence 108, App
21	1020	55.8	374	4 US-09-442-349A-6	Sequence 6, Appl
22	1020	55.8	374	4 US-09-442-349A-8	Sequence 8, Appl
23	1020	55.8	374	4 US-09-442-349A-10	Sequence 10, Appl
24	1020	55.8	374	4 US-09-442-349A-15	Sequence 15, Appl
25	1020	55.8	374	4 US-09-442-349A-17	Sequence 17, Appl
26	1020	55.8	374	4 US-09-442-349A-19	Sequence 19, Appl
27	1019	55.7	374	4 US-09-442-349A-1	Sequence 1, Appl

28	1019	55.7	374	4 US-09-442-349A-7	Sequence 7, Appl
29	1019	55.7	374	4 US-09-442-349A-9	Sequence 9, Appl
30	1019	55.7	374	4 US-09-442-349A-18	Sequence 18, Appl
31	1016	55.6	374	4 US-09-442-349A-11	Sequence 11, Appl
32	1016	55.6	374	4 US-09-442-349A-13	Sequence 13, Appl
33	1015	55.5	374	4 US-09-442-349A-5	Sequence 5, Appl
34	1015	55.5	374	4 US-09-442-349A-12	Sequence 12, Appl
35	1014	55.5	374	4 US-09-442-349A-36	Sequence 36, Appl
36	1014	55.5	374	4 US-09-442-349A-38	Sequence 38, Appl
37	1013	55.4	374	3 US-08-878-801-2	Sequence 2, Appl
38	1013	55.4	374	4 US-09-442-349A-3	Sequence 3, Appl
39	1013	55.4	374	4 US-09-442-349A-37	Sequence 37, Appl
40	1013	55.4	374	4 US-09-442-349A-48	Sequence 48, Appl
41	1013	55.4	374	4 US-09-442-349A-50	Sequence 50, Appl
42	1013	55.4	374	4 US-09-442-349A-54	Sequence 54, Appl
43	1013	55.4	374	4 US-09-442-349A-56	Sequence 56, Appl
44	1013	55.4	374	4 US-09-442-349A-105	Sequence 105, App
45	1012	55.4	374	4 US-09-442-349A-39	Sequence 39, Appl

ALIGNMENTS

RESULT 1
US-09-442-349A-29
; Sequence 29, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein chimera
US-09-442-349A-29

Query Match	57.0%;	Score 1042;	DB 4;	Length 374;			
Best Local Similarity	56.7%;	Pred. No. 1.2e-104;					
Matches	207;	Conservative	56;	Mismatches 88; Indels 14; Gaps 3;			
QY	4	CC--CLSAAEKESQRI	SAEIERHVRRDKDARRELKLL	LLGTGESGKSTFIKOMRIIHGS 61			
Db	9	CCPWCLTEDEKAAARVDQ	EINRILLEQKKQDRGELKLL	LLPGGESGKSTFIKOMRIIHGA 68			
QY	62	GYSDERKGF	TKLVYQNIFFAMQAMIRAMD	TLRIQYMCQNKENAQIIREVEVDKVTALS 121			
Db	69	GYSEERKGF	PLVYQNIFFVSMRAMIEMER	LQIPFSRPESKHASLVMSQDPYKVTTFE 128			
QY	122	RDQVAATKQLWLD	PGIQECYDRRREYQLSDS	AKYYLLTDIERIAMP	SFVPTQQDVLVRVVP 181		
Db	129	KRYAAAMQWLWRD	AGIRACYERRRRREFHLL	DSAVYYLSHLERITEEGYVPTAQDVLRSMP 188			
QY	182	TTGIIEY	PFLENIFRMVDVGGQSR	ERRKWIHCFESVTS	IIFLVALSEYDQVLAEC	DNE 241	
Db	189	TTGINEY	CFSVQKTNLRIVDVGGQ	SERKKWIHCFENV	IALIYLA	SLSEYDQCLEENNQE 248	
QY	242	NRMEESKAL	FRTIITYPWFNLSSV	ILFNKKDLEEKIM	YSHLISYFPEYTGPKQDV	KAA 301	
Db	249	NRMKESAL	EGTILELPWF	KSTSVILFNKTDILEE	KIPTSHLATYFP	SFGPKQDAEAA 308	
QY	302	RDFILKY	-----	QDQNPDK	E-KVIYSHFTCAT	DTENIRFVFAVKDTILQLNL 349	
Db	309	KRFILDMY	TRMYTGCVDG	PEGSNRN	KETKEIYSHFTCAT	DTKNIQVF	FEAVTVDVIIQNNL 368
QY	350	REFNL	354				

Db	129	KRYAAAMQWLWRDAGIRACYERREFHLLDSAVYYLSHLERITEEGYVPTAQDVLRSMP	188
Qy	182	TTGIIIEYFPFDLENIIFRMVDVGGQSRERRKWIHCFFESVTSIIFLVALSEYDQVLAEC	241
Db	189	TTGINEYCFSVQKTNLRIVDVGQKSERKKWIHCFFENVALIYLASLSEYDQCLEENNQE	248
Qy	242	NRMEESKALERTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFFPYTGPKQDV	301
Db	249	NRMKESALFEGTILELPWFKSTSVILFLNKTIDILEEKIPTSHLATYFSPFQGP	308
Qy	302	RDFILKLY-----QDQNPDK-KVIYSHFTCATDTENIRFVEAAVKDTTLQLNL	349
Db	309	KRFILDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTSNIQFVFEAVTDV	368
Qy	350	REFNL 354	
Db	369	KYIGL 373	

```

RESULT 5
US-09-442-349A-25
; Sequence 25, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 25
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
US-09-442-349A-25

```

Query Match	56.9%;	Score 1041;	DB 4;	Length 374;
Best Local Similarity	56.7%;	Pred. No. 1.5e-104;		
Matches 207; Conservative	55;	Mismatches 89;	Indels 14;	Gaps

Qy	4	CC--CLSAEESQRISAEIERHVRDCKDARRELKLLLTGTGESGKSTFIKOMRIIHGS	61
Dd	9	CCPWCLTEDEKAARVDQEINRILLEQQKKQDRGELKLLLPGSGSKSTFIKOMRIIHA	68
Qy	62	GYSDEDRKGFTKLVYQNIFTAQAMIRAMDTLRIQYMCEONKENAQIIREVEVDKVITALS	121
Dd	69	GYSEERKGFRLVYQNIFVSRAMIEAMERLQIPFSRPESKHASLVMSQDPYKVTFTE	128
Qy	122	RQVAAIKQLWLDPGIQECYDRREYQLSDSAKYLYLTDIERIA MPSFVPTQQDVLRVRVP	181
Dd	129	KRYAAMQWLWRDAGIRACYERRREFHLDSAVYYLSHLERITEEGYVPTAQDVLRSMP	188
Qy	182	TGTGIIEYPFDLENIIFRMVDVGQSRRKWIHCFSVTSIIFLVALSEYDQVLAECNE	241
Dd	189	TTGINEYCFSVQKTNLRIVDVGQKSERKKWIHC FENVIALIYLASLSYDQCLEENQE	248
Qy	242	NRMESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISIFYPEYTGPKQDKVAA	301
Dd	249	NRMKESALFGTILELPWFKSTSIVILFNKTDILEEKIPTSHLATYFPSPFQGPKQDAEA	308
Qy	302	RDFILKLY-----QDNPDKE-KVIYSHFTCATDTENIRVFVAVKDTILQLNL	349
Dd	309	KRFILDMYTRMYTCGVDPGCSNRNKETKEIYSHFTCATDTSNIQFVFQAVTOVIIQN	368
Qy	350	REFNL 354	
Dd	369	KYIGL 373	

```

RESULT 6
US-09-442-349A-30
; Sequence 30, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 30
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
; OTHER INFORMATION: chimera
US-09-442-349A-30

```

```

RESULT 7
US-09-442-349A-32
; Sequence 32, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein

```


QY 242 NRMEESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFFPEYTGPKQDVKAA 301
|||:| | | | | :| | | | | | | :| | | | | | | :| | | | | :| | | | |
Db 249 NRMKESLALFGTILELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFPFSGPKQDAEAA 308
|||:| | | | | :| | | | | | | :| | | | | | | :| | | | | :| | | | |
QY 302 RDFILKLY-----QDQNPDK-KVIYSHFTCATDTENIRFVFAAVKDTILQLNL 349
: | | | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 309 KRFIELDMYTRMYTGCVDGPEGSRNRNKETKEIYSHFTCATDTNNIQVFVNAVTDVLIQNNL 368
: | | | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 350 REFNL 354
: |
Db 369 KYIGL 373

RESULT 13
US-09-442-349A-26
; Sequence 26, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
; OTHER INFORMATION: chimera
US-09-442-349A-26

Query Match 56.7%; Score 1037; DB 4; Length 374;
Best Local Similarity 56.7%; Pred. No. 4.1e-104;
Matches 207; Conservative 55; Mismatches 89; Indels 14; Gaps 3;
QY 4 CC--CLSAEKEQSQRISAEIERHVRDKKARRELKLLLTGSGSGKSTFIKQMRIIHGS 61
|| | | : | | : | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db 9 CCPWCLTEDEKAAARVDQEI NRILLEQKKQDRGELKLLLPGESGSGKSTFIKQMRIIHGA 68
|||:| | | | | :| | | | | | | :| | | | | | | :| | | | | :| | | | |
QY 62 GYSDEDRKGFTKLVYQNIETAMQAMIRAMDTLRIQYMCQNKENAIIREVEVDKVTALS 121
|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 GYSEERKGFRLVYQNI FVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFE 128
|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 122 RDQVAAIKQLWLDPGIQECYDRRRREYQLSDSAKYIYLTDIERIA MPSFVPTQQDVLVRVP 181
: | | : | | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 129 KRYAAAMQWLWRDAGIRACYERRRREFHLLDSAVYYLSHLERITEEGYVPTAQDVLRSRMP 188
|||:| | | | | :| | | | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 182 TTGIIEYPPFDLENIIFRMVDVGGQSRERKRWIHCFSVTSIIFLVALSEYDQVLAECDNE 241
||| | | | : | : | | | | | :| | | | | | | :| | | | | | | :| | | | | | |
Db 189 TTGINEYCFSVQKTNLRIVDVGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQE 248
||| | | | : | : | | | | | :| | | | | | | :| | | | | | | :| | | | | | |
QY 242 NRMEESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPPEYTGPKQDVKAA 301
|||:| | | | | :| | | | | | | :| | | | | | | :| | | | | :| | | | |
Db 249 NRMKESLALFGTILELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFPFSGPKQDAEAA 308
|||:| | | | | :| | | | | | | :| | | | | | | :| | | | | :| | | | |
QY 302 RDFILKLY-----QDQNPDK-KVIYSHFTCATDTENIRFVFAAVKDTILQLNL 349
: | | | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 309 KRFIELDMYTRMYTGCVDGPEGSRNRNKETKEIYSHFTCATDTTCNIQVFVNAVTDVLIQNNL 368
: | | | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 350 REFNL 354
: |
Db 369 KYIGL 373

RESULT 14
US-09-442-349A-28
; Sequence 28, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:

; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
; OTHER INFORMATION: chimera
US-09-442-349A-28

Query Match 56.7%; Score 1037; DB 4; Length 374;
Best Local Similarity 56.7%; Pred. No. 4.1e-104;
Matches 207; Conservative 55; Mismatches 89; Indels 14; Gaps 3;
QY 4 CC--CLSAEKEQSQRISAEIERHVRDKKARRELKLLLTGSGSGKSTFIKQMRIIHGS 61
|| | | : | | : | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db 9 CCPWCLTEDEKAAARVDQEI NRILLEQKKQDRGELKLLLPGESGSGKSTFIKQMRIIHGA 68
|||:| | | | | :| | | | | | | :| | | | | | | :| | | | | :| | | | |
QY 62 GYSDEDRKGFTKLVYQNIETAMQAMIRAMDTLRIQYMCQNKENAIIREVEVDKVTALS 121
|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 GYSEERKGFRLVYQNI FVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFE 128
|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 122 RDQVAAIKQLWLDPGIQECYDRRRREYQLSDSAKYIYLTDIERIA MPSFVPTQQDVLVRVP 181
: | | : | | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 129 KRYAAAMQWLWRDAGIRACYERRRREFHLLDSAVYYLSHLERITEEGYVPTAQDVLRSRMP 188
|||:| | | | | :| | | | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 182 TTGIIEYPPFDLENIIFRMVDVGGQSRERKRWIHCFSVTSIIFLVALSEYDQVLAECDNE 241
||| | | | : | : | | | | | :| | | | | | | :| | | | | | | :| | | | | | |
Db 189 TTGINEYCFSVQKTNLRIVDVGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQE 248
||| | | | : | : | | | | | :| | | | | | | :| | | | | | | :| | | | | | |
QY 242 NRMEESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPPEYTGPKQDVKAA 301
|||:| | | | | :| | | | | | | :| | | | | | | :| | | | | :| | | | |
Db 249 NRMKESLALFGTILELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFPFSGPKQDAEAA 308
|||:| | | | | :| | | | | | | :| | | | | | | :| | | | | :| | | | |
QY 302 RDFILKLY-----QDQNPDK-KVIYSHFTCATDTENIRFVFAAVKDTILQLNL 349
: | | | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 309 KRFIELDMYTRMYTGCVDGPEGSRNRNKETKEIYSHFTCATDTTCNIQVFVNAVTDVLIQNNL 368
: | | | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 350 REFNL 354
: |
Db 369 KYIGL 373

RESULT 15
US-09-442-349A-20
; Sequence 20, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
; OTHER INFORMATION: chimera
US-09-442-349A-20

Query Match 56.7%; Score 1036; DB 4; Length 374;
Best Local Similarity 56.7%; Pred. No. 5.3e-104;
Matches 207; Conservative 55; Mismatches 89; Indels 14; Gaps 3;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 14:11:47 ; Search time 13 Seconds
(without alignments)
551.030 Million cell updates/sec

Title: US-09-492-028-2
Perfect score: 1828
Sequence: 1 MAGCCCLSAEKEQSORISAE.....VFAAVKDTILQLNLFNLY 355

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1537	84.1	353	10 US-09-984-292-8	Sequence 8, Appli
2	1537	84.1	353	10 US-09-989-497-8	Sequence 8, Appli
3	1537	84.1	359	10 US-09-984-292-9	Sequence 9, Appli
4	1537	84.1	359	10 US-09-989-497-9	Sequence 9, Appli
5	1535	84.0	359	10 US-09-989-295-2	Sequence 2, Appli
6	1530	83.7	359	10 US-09-984-292-11	Sequence 11, Appl
7	1530	83.7	359	10 US-09-989-497-11	Sequence 11, Appl
8	1527	83.5	353	10 US-09-984-292-2	Sequence 2, Appli
9	1527	83.5	353	10 US-09-984-292-17	Sequence 17, Appl
10	1527	83.5	353	10 US-09-989-497-2	Sequence 2, Appli
11	1527	83.5	353	10 US-09-989-497-17	Sequence 17, Appl
12	1527	83.5	359	10 US-09-984-292-1	Sequence 1, Appli
13	1527	83.5	359	10 US-09-984-292-21	Sequence 21, Appl
14	1527	83.5	359	10 US-09-989-497-1	Sequence 1, Appli
15	1527	83.5	359	10 US-09-989-497-21	Sequence 21, Appl
16	1526	83.5	353	10 US-09-984-292-4	Sequence 4, Appli
17	1526	83.5	353	10 US-09-984-292-5	Sequence 5, Appli
18	1526	83.5	353	10 US-09-989-497-4	Sequence 4, Appli
19	1526	83.5	353	10 US-09-989-497-5	Sequence 5, Appli

20	1526	83.5	359	10 US-09-984-292-3	Sequence 3, Appli
21	1526	83.5	359	10 US-09-989-497-3	Sequence 3, Appli
22	1522	83.3	359	10 US-09-984-292-26	Sequence 26, Appl
23	1522	83.3	359	10 US-09-989-497-26	Sequence 26, Appl
24	1520	83.2	353	10 US-09-899-295-4	Sequence 4, Appli
25	1520	83.2	353	10 US-09-899-295-6	Sequence 6, Appli
26	1520	83.2	353	10 US-09-899-295-8	Sequence 8, Appli
27	1519	83.1	353	10 US-09-984-292-6	Sequence 6, Appli
28	1519	83.1	353	10 US-09-984-292-16	Sequence 16, Appl
29	1519	83.1	353	10 US-09-989-497-6	Sequence 6, Appli
30	1519	83.1	353	10 US-09-989-497-16	Sequence 16, Appl
31	1519	83.1	359	10 US-09-984-292-10	Sequence 10, Appl
32	1519	83.1	359	10 US-09-984-292-12	Sequence 12, Appl
33	1519	83.1	359	10 US-09-989-497-10	Sequence 10, Appl
34	1519	83.1	359	10 US-09-989-497-12	Sequence 12, Appl
35	1518	83.0	353	10 US-09-984-292-14	Sequence 14, Appl
36	1518	83.0	353	10 US-09-989-497-14	Sequence 14, Appl
37	1517	83.0	359	10 US-09-984-292-37	Sequence 37, Appl
38	1517	83.0	359	10 US-09-989-497-37	Sequence 37, Appl
39	1516	82.9	359	10 US-09-984-292-20	Sequence 20, Appl
40	1516	82.9	359	10 US-09-984-292-22	Sequence 22, Appl
41	1516	82.9	359	10 US-09-989-497-20	Sequence 20, Appl
42	1516	82.9	359	10 US-09-989-497-22	Sequence 22, Appl
43	1515	82.9	353	10 US-09-984-292-24	Sequence 24, Appl
44	1515	82.9	353	10 US-09-989-497-24	Sequence 24, Appl
45	1515	82.9	359	10 US-09-984-292-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-09-984-292-8
; Sequence 8, Application US/09984292
; Patent No. US20020128433A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF GENERAL INFORMATION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280649
; CURRENT APPLICATION NUMBER: US/09/984,292
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-984-292-8

Query Match	84.1%	Score	1537;	DB	10;	Length	353;
Best Local Similarity	82.3%	Pred. No.	2e-131;				
Matches	289;	Conservative	34;	Mismatches	28;	Indels	0;
Gaps	0;						
Qy	5	CCLSAEKEQSQRISAEIERHVRDRKDDARRELKLLLGTS	EGSGKSTFIKQMRIHSGYS	64			
Db	3	CCLSEEAKEARRINDEIERHVRDRKDDARRELKLLLGTS	EGSGKSTFIKQMRIHSGYS	62			
Qy	65	DEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ	124				
Db	63	DEDRKGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLREVVDVEKVSADFVDP	122				
Qy	125	VAAIKQLWLDPGIQECYDRRRREYQLSDSAKYLLTDIERIAMP	SFVPTQQDVLVRVPTTG	184			
Db	123	VAAIKSLWNDPGIQECYDRRRREYQLSDSTKYKYYLNDLDRVADPSY	LPTQQDVLVRVPTTG	182			
Qy	185	IIEYFPDLNIIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDNENRM	244				
Db	183	IIEYFPDLQSVIFRMVDVGGQSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM	242				

```
QY 245 EESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFFPYTGPKQDVKAARDF 304
|||||
Db 243 EESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFFPYDGPQRDAQAAREF 302
|||||
QY 305 ILKLYQDQNPDKVKIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
|||||
Db 303 ILKMFVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 353

RESULT 2
US-09-989-497-8
; Sequence 8, Application US/09989497
; Patent No. US20020143151A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280735
; CURRENT APPLICATION NUMBER: US/09/989,497
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/984,292
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-989-497-8

Query Match 84.1%; Score 1537; DB 10; Length 353;
Best Local Similarity 82.3%; Pred. No. 2e-131;
Matches 289; Conservative 34; Mismatches 28; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRIISAEIERHVRDRKDDARRELKLLLLGTGSGKSTFIKQMRIIHGSGYS 64
|||||
Db 3 CCLSEEAKEARRINDEIERHVRDRKDDARRELKLLLLGTGSGKSTFIKQMRIIHGSGYS 62
|||||

QY 65 DEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ 124
|||||
Db 63 DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSADFVDP 122
|||||

QY 125 VAAIKQLWDPGIECYDRRREYQLSDSAKYIYLTDIERIA MPSFVPTQQDVLVRVPTTG 184
|||||
Db 123 YAAIKSLWNDPGIECYDRRREYQLSDSTKYIYLNLDLDRVADPSYLP TQQDVLVRVPTTG 182
|||||

QY 185 IIEYFPDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDNENRM 244
|||||
Db 183 IIEYFPDLQSVIFRMVDVGGQSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 242
|||||

QY 245 EESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFFPYTGPKQDVKAARDF 304
|||||
Db 243 EESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFFPYDGPQRDAQAAREF 302
|||||

QY 305 ILKLYQDQNPDKVKIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
|||||
Db 303 ILKMFVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 353

RESULT 3
US-09-984-292-9
; Sequence 9, Application US/09984292
; Patent No. US20020128433A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
```

```
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280649
; CURRENT APPLICATION NUMBER: US/09/984,292
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-984-292-9

Query Match 84.1%; Score 1537; DB 10; Length 359;
Best Local Similarity 82.3%; Pred. No. 2e-131;
Matches 289; Conservative 34; Mismatches 28; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRIISAEIERHVRDRKDDARRELKLLLLGTGSGKSTFIKQMRIIHGSGYS 64
|||||
Db 9 CCLSEEAKEARRINDEIERHVRDRKDDARRELKLLLLGTGSGKSTFIKQMRIIHGSGYS 68
|||||

QY 65 DEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ 124
|||||
Db 69 DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSADFVDP 128
|||||

QY 125 VAAIKQLWDPGIECYDRRREYQLSDSAKYIYLTDIERIA MPSFVPTQQDVLVRVPTTG 184
|||||
Db 129 YAAIKSLWNDPGIECYDRRREYQLSDSTKYIYLNLDLDRVADPSYLP TQQDVLVRVPTTG 188
|||||

QY 185 IIEYFPDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDNENRM 244
|||||
Db 189 IIEYFPDLQSVIFRMVDVGGQSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248
|||||

QY 245 EESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFFPYTGPKQDVKAARDF 304
|||||
Db 249 EESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFFPYDGPQRDAQAAREF 308
|||||

QY 305 ILKLYQDQNPDKVKIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
|||||
Db 309 ILKMFVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
|||||

RESULT 4
US-09-989-497-9
; Sequence 9, Application US/09989497
; Patent No. US20020143151A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280735
; CURRENT APPLICATION NUMBER: US/09/989,497
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/984,292
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-989-497-9

Query Match 84.1%; Score 1537; DB 10; Length 359;
Best Local Similarity 82.3%; Pred. No. 2e-131;
Matches 289; Conservative 34; Mismatches 28; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRIISAEIERHVRDRKDDARRELKLLLLGTGSGKSTFIKQMRIIHGSGYS 64
```


Db 9 CCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLTGTGESGKSTFIQMRIIHGSGYS 68
QY 65 DEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ 124
Db 69 DEDRKGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSADFVPD 128
QY 125 VAAIKQLWLDPGIQECYDRRREYQLSDSAKYIYLTDIERIAMPSFVPTQDDVLRVRVPTTG 184
Db 129 YAAIKSLWNDDPGIQECYDRRREYQLSDSTKYIYNLDRVADPSYLTQDDVLRVRVPTTG 188
QY 185 IIEYPPFDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIIFLVALSEYDQVLAECDNENRM 244
Db 189 IIEYPPFDLQSVIFRMVDVGGQSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248
QY 245 EESKALFRITIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAARDF 304
Db 249 EESKALFRITIITYPWFQNSVILFLNKKDLLEEKIMYSHLVDYFFPEYDGPQDAQAAREF 308
QY 305 ILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
Db 309 ILKMFVDLNPDSDKIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 5
US-09-899-295-2
; Sequence 2, Application US/09899295
; Patent No. US20020127601A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Pharma Deutschland GmbH
; TITLE OF INVENTION: Process for identifying modulators of G protein coupled
; TITLE OF INVENTION: receptors
; FILE REFERENCE: AVE D-2000/A033 englisch
; CURRENT APPLICATION NUMBER: US/09/899,295
; CURRENT FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-899-295-2

Query Match 84.0%; Score 1535; DB 10; Length 359;
Best Local Similarity 82.1%; Pred. No. 3.1e-131;
Matches 288; Conservative 35; Mismatches 28; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRIISAEIERHVRRDKRDARRELKLLLTGTGESGKSTFIQMRIIHGSGYS 64
Db 9 CCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLTGTGESGKSTFIQMRIIHGSGYS 68
QY 65 DEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ 124
Db 69 DEDRKGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPY 128
QY 125 VAAIKQLWLDPGIQECYDRRREYQLSDSAKYIYLTDIERIAMPSFVPTQDDVLRVRVPTTG 184
Db 129 VDAIKSLWNDDPGIQECYDRRREYQLSDSTKYIYNLDRVADPAYLPTQDDVLRVRVPTTG 188
QY 185 IIEYPPFDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIIFLVALSEYDQVLAECDNENRM 244
Db 189 IIEYPPFDLQSVIFRMVDVGGQSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248
QY 245 EESKALFRITIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAARDF 304
Db 249 EESKALFRITIITYPWFQNSVILFLNKKDLLEEKIMYSHLVDYFFPEYDGPQDAQAAREF 308
QY 305 ILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
Db 309 ILKMFVDLNPDSDKIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 6

US-09-984-292-11
; Sequence 11, Application US/09984292
; Patent No. US20020128433A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280649
; CURRENT APPLICATION NUMBER: US/09/984,292
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-984-292-11

Query Match 83.7%; .Score 1530; DB 10; Length 359;
Best Local Similarity 82.1%; Pred. No. 8.7e-131;
Matches 288; Conservative 34; Mismatches 29; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRIISAEIERHVRRDKRDARRELKLLLTGTGESGKSTFIQMRIIHGSGYS 64
Db 9 CCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLTGTGESGKSTFIQMRIIHGSDYS 68
QY 65 DEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ 124
Db 69 DEDRKGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSADFVPD 128
QY 125 VAAIKQLWLDPGIQECYDRRREYQLSDSAKYIYLTDIERIAMPSFVPTQDDVLRVRVPTTG 184
Db 129 YAAIKSLWNDDPGIQECYDRRREYQLSDSTKYIYNLDRVADPSYLTQDDVLRVRVPTTG 188
QY 185 IIEYPPFDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIIFLVALSEYDQVLAECDNENRM 244
Db 189 IIEYPPFDLQSVIFRMVDVGGQSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248
QY 245 EESKALFRITIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAARDF 304
Db 249 EESKALFRITIITYPWFQNSVILFLNKKDLLEEKIMYSHLVDYFFPEYDGPQDAQAAREF 308
QY 305 ILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
Db 309 ILKMFVDLNPDSDKIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 7
US-09-989-497-11
; Sequence 11, Application US/09989497
; Patent No. US20020143151A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280735
; CURRENT APPLICATION NUMBER: US/09/989,497
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/984,292
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 359
; TYPE: PRT

Qy 125 VAAIKQLWLDPGIQECYDRRREYQLSDSAKYLLTDIERIAMPSFVPTQQDVLVRVPPTG 184
Db 129 YAAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLVRVPPTG 188
Qy 185 IIEYPFDLENIFIRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDNENRM 244
Db 189 IIEYPFDLQSVIFIRMVDVGGQSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248
Qy 245 EESKALFRITITYPWFLNSSVILFLNKKDLLEEKIMYSHLISYFFEYTGPKQDVKAARDF 304
Db 249 EESKALFRITITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDPFEYDGPQORDAQAAREF 308
Qy 305 ILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
Db 309 ILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

Search completed: February 3, 2003, 14:15:53
Job time : 14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 14:08:56 ; Search time 13 Seconds
(without alignments)
1132.623 Million cell updates/sec

Title: US-09-492-028-2
Perfect score: 1828
Sequence: 1 MAGCCCLSAEKEQSRIAE.....VFAAVKDTILQLNLFNVLV 355

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1828	100.0	355	1	GB14_MOUSE	P30677 mus musculu
2	1780	97.4	355	1	GB14_BOVIN	P38408 bos taurus
3	1771	96.9	355	1	GB14_HUMAN	O95837 homo sapien
4	1643.5	89.9	354	1	GB14_XENLA	O73819 xenopus lae
5	1538	84.1	353	1	GBQ_MOUSE	P21279 mus musculu
6	1527	83.5	359	1	GB11_MEIGA	P45645 meleagris g
7	1526	83.5	353	1	GBQ_RAT	P82471 rattus norv
8	1524	83.4	353	1	GBQ_CANFA	Q28294 canis famil
9	1524	83.4	359	1	GB11_HUMAN	P29992 homo sapien
10	1519	83.1	353	1	GBQ_HUMAN	P50148 homo sapien
11	1518	83.0	359	1	GB11_BOVIN	P38409 bos taurus
12	1517	83.0	359	1	GB11_XENLA	P43444 xenopus lae
13	1512	82.7	359	1	GB11_MOUSE	P21278 mus musculu
14	1510	82.6	359	1	GB11_RAT	Q9jid2 rattus norv
15	1492	81.6	353	1	GBQ_XENLA	P38410 xenopus lae
16	1416	77.5	353	1	GBQ_PATYE	O15975 patinopecte
17	1407	77.0	353	1	GBQ_HOMAM	P91950 homarus ame
18	1406	76.9	353	1	GBQ_LYMST	P38411 lymnaea sta
19	1382.5	75.6	354	1	GBQ_LOLFO	P38412 loligo forb
20	1346	73.6	353	1	GBQ_DROME	P23625 drosophila
21	1022	55.9	374	1	GB15_MOUSE	P30678 mus musculu
22	1013	55.4	374	1	GB15_HUMAN	P30679 homo sapien
23	1002	54.8	374	1	GB15_RAT	O88302 rattus norv
24	910.5	49.8	353	1	GB02_HUMAN	P29777 homo sapien
25	905.5	49.5	353	1	GB02_CRILO	P17806 cricetulus
26	900.5	49.3	353	1	GB0_XENLA	P10825 xenopus lae
27	899.5	49.2	353	1	GB02_MOUSE	P18873 mus musculu
28	899.5	49.2	353	1	GB0_LYMST	P30683 lymnaea sta
29	898.5	49.2	353	1	GB0_HELTI	P51877 helisoma tr
30	895.5	49.0	354	1	GB01_DROME	P16378 drosophila
31	895	49.0	354	1	GBI2_CHICK	P50147 gallus gall
32	894.5	48.9	353	1	GB02_RAT	P30033 rattus norv
33	894	48.9	354	1	GBI2_RAT	P04897 rattus norv

34	893	48.9	354	1	GBI2_CANFA	P38400 canis faml1
35	893	48.9	354	1	GBI2_MOUSE	P08752 mus musculu
36	892.5	48.8	353	1	GBI1_HUMAN	P04898 homo sapien
37	892.5	48.8	353	1	GBI1_ASTPE	P30676 asterina pe
38	891.5	48.8	353	1	GBI1_RAT	P10824 rattus norv
39	891	48.7	354	1	GBI2_HUMAN	P04899 homo sapien
40	891	48.7	354	1	GBI2_ORYLA	O13055 oryzias lat
41	890.5	48.7	353	1	GBI1_CHICK	P50146 gallus gall
42	889.5	48.7	353	1	GBI1_CAVPO	P38401 cavia porce
43	889	48.6	354	1	GBI2_CAVPO	P38402 cavia porce
44	888	48.6	356	1	GB0_PATYE	O15976 patinopecte
45	884.5	48.4	353	1	GBI1_XENLA	P27044 xenopus lae

ALIGNMENTS

RESULT 1

GB14_MOUSE

ID	GB14_MOUSE	STANDARD;	PRT;	355 AA.
AC	P30677;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Guanine nucleotide-binding protein, alpha-14 subunit.			
GN	GNA14 OR GNA-14.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92052208; PubMed=1946421;			
RA	Wilkie T.M., Scherle P.A., Strathmann M.P., Slepak V.Z., Simon M.I.;			
RT	"Characterization of G-protein alpha subunits in the Gq class: expression in murine tissues and in stromal and hematopoietic cell lines.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:10049-10053(1991).			
RN	[2]			
RP	SEQUENCE OF 217-267 FROM N.A.			
RX	MEDLINE=90017488; PubMed=2508088;			
RA	Strathmann M., Wilkie T.M., Simon M.I.;			
RT	"Diversity of the G-protein family: sequences from five additional alpha subunits in the mouse.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:7407-7409(1989).			
CC	-!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE SIGNALING SYSTEMS.			
CC	-!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA). THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.			
CC	-!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M80631; AAA83222.1; -.			
DR	EMBL; M57616; AAA63304.1; -.			
DR	PIR; A41534; A41534.			
DR	PIR; E33833; E33833.			
DR	HSSP; P10824; IBOF.			
DR	MGD; MGI:95769; Gna14.			
DR	InterPro; IPR001019; Gprotein_alpha.			
DR	Pfam; PF00503; G-alpha; 1.			
DR	PRINTS; PR00318; GPROTEINA.			
DR	ProDom; PD000281; Gprotein_alpha; 1.			
DR	SMART; SM00275; G-alpha; 1.			
KW	GTP-binding; Transducer; Multigene family; ADP-ribosylation.			
FT	NP_BIND 42 49 GTP (BY SIMILARITY).			

```
FT NP_BIND 201 205 GTP (BY SIMILARITY).
FT NP_BIND 270 273 GTP (BY SIMILARITY).
FT MOD_RES 179 179 ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT (BY SIMILARITY).
SQ SEQUENCE 355 AA; 41522 MW; A4810D72169878E9 CRC64;

Query Match 100.0%; Score 1828; DB 1; Length 355;
Best Local Similarity 100.0%; Pred. No. 6.5e-130;
Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGCCCLSAEKEESQRISAEIERHVRDKKDDARRELKLLLTGTGSGKSTFIKQMRIHG 60
   |||||
Db 1 MAGCCCLSAEKEESQRISAEIERHVRDKKDDARRELKLLLTGTGSGKSTFIKQMRIHG 60

QY 61 SGYSDEDRKGFTKLVIYNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTAL 120
   |||||
Db 61 SGYSDEDRKGFTKLVIYNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTAL 120

QY 121 SRDQVAAIKQLWLDPGIQECYDRRREYQLSDSAKYLYLTDIRIAMPSEFVPTQQDVLVRV 180
   |||||
Db 121 SRDQVAAIKQLWLDPGIQECYDRRREYQLSDSAKYLYLTDIRIAMPSEFVPTQQDVLVRV 180

QY 181 PTTGIIEYPFDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDN 240
   |||||
Db 181 PTTGIIEYPFDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDN 240

QY 241 ENRMEESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKA 300
   |||||
Db 241 ENRMEESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKA 300

QY 301 ARDFILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
   |||||
Db 301 ARDFILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355

RESULT 2
GB14_BOVIN
ID GB14_BOVIN STANDARD; PRT; 355 AA.
AC P38408;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanine nucleotide-binding protein, alpha-14 subunit (GL1).
GN GNA14.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91286303; PubMed=1905731;
RA Nakamura F., Ogata K., Shiozaki K., Kameyama K., Ohara K., Haga T.,
RA Nukada T.;
RT "Identification of two novel GTP-binding protein alpha-subunits that
RT lack apparent ADP-ribosylation sites for pertussis toxin.";
RL J. Biol. Chem. 266:12676-12681(1991).
CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; D90335; BAA14349.1; -.
```

```
DR PIR; A40891; A40891.
DR HSSP; P10824; 1BOF.
DR InterPro; IPR001019; Gprotein_alpha.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR ProDom; PD000281; Gprotein_alpha; 1.
DR SMART; SM00275; G-alpha; 1.
KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT NP_BIND 42 49 GTP (BY SIMILARITY).
FT NP_BIND 201 205 GTP (BY SIMILARITY).
FT NP_BIND 270 273 GTP (BY SIMILARITY).
FT MOD_RES 179 179 ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT (BY SIMILARITY).
SQ SEQUENCE 355 AA; 41498 MW; 60888CC6C9B7243B CRC64;

Query Match 97.4%; Score 1780; DB 1; Length 355;
Best Local Similarity 96.6%; Pred. No. 2.6e-126;
Matches 343; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAGCCCLSAEKEESQRISAEIERHVRDKKDDARRELKLLLTGTGSGKSTFIKQMRIHG 60
   |||||
Db 1 MAGCCCLSAEKEESQRISAEIERHVRDKKDDARRELKLLLTGTGSGKSTFIKQMRIHG 60

QY 61 SGYSDEDRKGFTKLVIYNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTAL 120
   |||||
Db 61 SGYSDEDRKGFTKLVIYNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTAL 120

QY 121 SRDQVAAIKQLWLDPGIQECYDRRREYQLSDSAKYLYLTDIRIAMPSEFVPTQQDVLVRV 180
   |||||
Db 121 SRDQVAAIKQLWLDPGIQECYDRRREYQLSDSAKYLYLTDIRIAMPSEFVPTQQDVLVRV 180

QY 181 PTTGIIEYPFDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDN 240
   |||||
Db 181 PTTGIIEYPFDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDN 240

QY 241 ENRMEESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKA 300
   |||||
Db 241 ENRMEESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKA 300

QY 301 ARDFILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
   |||||
Db 301 ARDFILKLYQDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355

RESULT 3
GB14_HUMAN
ID GB14_HUMAN STANDARD; PRT; 355 AA.
AC Q95837;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanine nucleotide-binding protein, alpha-14 subunit (G-protein alpha
DE subunit 14).
GN GNA14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99208664; PubMed=10191087;
RA Rubio J.P., Levy E.R., Dobson-Stone C., Monaco A.P.;
RT "Genomic organization of the human G-alpha-14 and G-alpha-Q genes and
RT mutation analysis in chorea-acanthocytosis (CHAC).";
RL Genomics 57:84-93(1999).
CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```


Db 129 VSAIKTLWNDPGIECYDRRRREYQLSDSAKYILSDVDRIATPGYLPQTQDVLVRVPTTG 188

QY 185 IIEYFPDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDNENRM 244

Db 189 IIEYFPDLENIIFRMVDVGGQSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248

QY 245 EESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAARDF 304

Db 249 EESKALFRTIITYPWFQNSVILFLNKKDLLEDKILYSHLVDYFPEFDGPDQDAQAAREF 308

QY 305 ILKLYQDQNPDKKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355

Db 309 ILKMEVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 7

GBQ_RAT

ID GBQ_RAT STANDARD; PRT; 353 AA.

AC P82471;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Guanine nucleotide-binding protein G(q), alpha subunit.

GN GNAQ.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Strotmann R.;

RT "Rat G alpha q subunit.";

RL Submitted (FEB-2000) to the EMBL/GenBank/DBBJ databases.

RN [2]

RP SEQUENCE OF 74-229 FROM N.A.

RC TISSUE=Brain cortex;

RX MEDLINE=93081611; PubMed=1333286;

RA Lesch K.-P., Manji H.K.;

RT "Signal-transducing G proteins and antidepressant drugs: evidence for

RT modulation of alpha subunit gene expression in rat brain.";

RL Biol. Psychiatry 32:549-579(1992).

RN [3]

RP SEQUENCE OF 238-331 FROM N.A.

RC STRAIN=Sprague-Dawley; TISSUE=Kidney cortex;

RA Thomas C.P.;

RT "GTP-binding protein expression in glomerular mesangial cells.";

RL Submitted (JAN-1995) to the EMBL/GenBank/DBBJ databases.

CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE SIGNALING SYSTEMS.

CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).

CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.

CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF234260; AAF59930.1; ALT_INIT.

DR EMBL: L37294; AAB02848.1; -.

DR HSSP: P10824; IBOF.

DR InterPro: IPR001019; Gprotein_alpha.

DR Pfam: PF00503; G-alpha; 1.

DR ProDom: PD000281; Gprotein_alpha; 1.

DR SMART: SM00275; G-alpha; 1.

KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;

KW Palmitate; Lipoprotein.

FT LIPID 3 PALMITATE (BY SIMILARITY).

FT LIPID 4 4 PALMITATE (BY SIMILARITY).

FT NP_BIND 40 47 GTP (BY SIMILARITY).

FT NP_BIND 199 203 GTP (BY SIMILARITY).

FT NP_BIND 268 271 GTP (BY SIMILARITY).

FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX)

FT (BY SIMILARITY).

FT CONFLICT 85 85 V -> I (IN REF. 2).

SQ SEQUENCE 353 AA; 41469 MW; BB4C211FD4DD47534 CRC64;

Query Match 83.5%; Score 1526; DB 1; Length 353;

Best Local Similarity 81.5%; Pred. NO. 2.7e-107;

Matches 286; Conservative 36; Mismatches 29; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRIISAEIERHVRDKKDDARRELKLLLTGSGSGKSTFIKQMRIIHGSGYS 64

Db 3 CCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLTGSGSGKSTFIKQMRIIHGSGYS 62

QY 65 DEDRKGFTKLVIYQNIIFTAMQAMIRAMDTLRIQYMCQENKENAQIIREVEVDKVTALSRDQ 124

Db 63 DEDKRGFTKLVIYQNIIFTAMQAMVIRAMDTLKIPIKYEHKHAQALVREVDVEKVSAPENPY 122

QY 125 VAAIKQLWLDPGIECYDRRREYQLSDSAKYIYLTDIERIAMPSFVPTQDVLVRVPTTG 184

Db 123 VDAIKSLWNDPGIECYDRRREYQLSDSTKYIYLNLDLRVADPSYLPQTQDVLVRVPTTG 182

QY 185 IIEYFPDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDNENRM 244

Db 183 IIEYFPDLQSVIFRMVDVGGQSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 242

QY 245 EESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAARDF 304

Db 243 EESKALFRTIITYPWFQNSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQDAQAAREF 302

QY 305 ILKLYQDQNPDKKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355

Db 303 ILKMEVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 353

RESULT 8

GBQ_CANFA

ID GBQ_CANFA STANDARD; PRT; 353 AA.

AC Q28294;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Guanine nucleotide-binding protein G(q), alpha subunit.

GN GNAQ.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxID=9615;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96433124; PubMed=8836152;

RA Johnson G.J., Leis L.A., Dunlop P.C.;

RT "Specificity of G alpha q and G alpha 11 gene expression in platelets

RT and erythrocytes. Expressions of cellular differentiation and species

RT differences.";

RL Biochem. J. 318:1023-1031(1996).

CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE SIGNALING SYSTEMS.

CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).

CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.

CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----


```
CC -----
DR EMBL; L76257; AAB05548.1; ALT_INIT.
DR HSSP; P10824; IBOF.
DR InterPro; IPR001019; Gprotein_alpha.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR ProDom; PD000281; Gprotein_alpha; 1.
DR SMART; SM00275; G-alpha; 1.
KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW Palmitate; Lipoprotein.
FT LIPID 3 3 PALMITATE (BY SIMILARITY).
FT LIPID 4 4 PALMITATE (BY SIMILARITY).
FT NP_BIND 40 47 GTP (BY SIMILARITY).
FT NP_BIND 199 203 GTP (BY SIMILARITY).
FT NP_BIND 268 271 GTP (BY SIMILARITY).
FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT (BY SIMILARITY).
SQ SEQUENCE 353 AA; 41467 MW; E7737E07B1F4904C CRC64;

Query Match 83.4%; Score 1524; DB 1; Length 353;
Best Local Similarity 81.5%; Pred. No. 3.8e-107;
Matches 286; Conservative 36; Mismatches 29; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRISAEIERHVRDCKDARRELKLLLLGTGESGKSTFIKQMRIHSGYS 64
Db 3 CCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIHSGYS 62

QY 65 DEDRKGFTKLVIYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ 124
Db 63 DEDKRGFTKLVIYQNIFTAMQAMIRAMDTLKIPIKYEHNKHAHAQLVREVDVEKVSAFENPY 122

QY 125 VAAIKQLWLDPGIECYDRRREYQLSDSAKYVLTDIERIAMPSFVPTQQDVLVRVPTTG 184
Db 123 VDAIKSLWNDPGIECYDRRREYQLSDSTKYVLLNLDLRVADPAYLPTQQDVLVRVPTTG 182

QY 185 IIEYFDFLENIIFRMVDVGGQSRERRKWIHCFSVTSIIFLVALSEYDQVLAECDNENRM 244
Db 183 IIEYFDFLQSVIFRMVDVGGQSRERRKWIHCFNVTSIMFLVALSEYDQVLVESDNENRM 242

QY 245 EESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAARDF 304
Db 243 EESKALFRTIITYPWFQNSVILFLNKKDLLEEKIMYSHLVDFPEYDGPORDAQAAREF 302

QY 305 ILKLYQDQNPDKVKIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
Db 303 ILKMFVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 353

RESULT 9
GB11_HUMAN STANDARD; PRT; 359 AA.
AC P29992; Q14350; O15109;
DT 01-APR-1993 (Rel. 25, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanine nucleotide-binding protein G(Y), alpha subunit (Alpha-11).
GN GN11 OR G11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=91219481; PubMed=1902575;
RA Jiang M., Pandey S., Tran V.T., Fong H.K.W.;
RT "Guanine nucleotide-binding regulatory proteins in retinal pigment
RT epithelial cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:3907-3911(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Bai X.H., Acharya R., Bai Y.H., Murtagh J.J.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBDJ databases.
```

```

[3]
RN SEQUENCE FROM N.A.
RP Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A., Montgomey M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBDJ databases.
[4]
RN SEQUENCE OF 244-337 FROM N.A.
RP TISSUE=Hematopoietic;
RC MEDLINE=96077138; PubMed=7492305;
RX Thomas C.P., Dunn M.J., Mattera R.;
RT "Ca2+ signalling in K562 human erythroleukaemia cells: effect of
RT dimethyl sulphoxide and role of G-proteins in thrombin- and
RT thromboxane A2-activated pathways.";
RL Biochem. J. 312:151-158(1995).
CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M69013; AAA58624.1; -.
DR EMBL; AF011497; AAB64303.1; -.
DR EMBL; AC005262; AAC25615.1; -.
DR EMBL; L40630; AAA99949.1; -.
DR PIR; A39394; RGHUGY.
DR HSSP; P10824; IBOF.
DR Genew; HGNC:4379; GN11.
DR MIM; 139313; -.
DR InterPro; IPR001019; Gprotein_alpha.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR ProDom; PD000281; Gprotein_alpha; 1.
DR SMART; SM00275; G-alpha; 1.
KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT NP_BIND 46 53 GTP (BY SIMILARITY).
FT NP_BIND 205 209 GTP (BY SIMILARITY).
FT NP_BIND 274 277 GTP (BY SIMILARITY).
FT MOD_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT (BY SIMILARITY).
FT CONFLICT 6 6 M -> I (IN REF. 2).
FT CONFLICT 266 266 N -> H (IN REF. 4).
FT CONFLICT 285 285 Y -> H (IN REF. 4).
FT CONFLICT 301 302 DA -> EP (IN REF. 1).
FT CONFLICT 310 310 L -> P (IN REF. 2).
SQ SEQUENCE 359 AA; 42123 MW; DD37176589E66046 CRC64;

Query Match 83.4%; Score 1524; DB 1; Length 359;
Best Local Similarity 82.1%; Pred. No. 3.9e-107;
Matches 288; Conservative 35; Mismatches 28; Indels 0; Gaps 0;
```

```

QY 5 CCLSAEEKESQRISAEIERHVRDCKDARRELKLLLLGTGESGKSTFIKQMRIHSGYS 64
Db 9 CCLSDVEKESKRINAEIEKQLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIHSGYS 68

QY 65 DEDRKGFTKLVIYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ 124
Db 69 EEDKRGFTKLVIYQNIFTAMQAMIRAMETLKILYKYEQNKANALLIREVDVEKVTTFEHQY 128
```



```
RT "Rattus norvegicus guanine nucleotide binding protein alpha 11 subunit
RL (G11).";
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF239674; AAF81690.1; -
CC HSSP; P10824; IBOF.
CC InterPro; IPR001019; Gprotein_alpha.
CC Pfam; PF00503; G-alpha; 1.
CC PRINTS; PR00318; GPROTEINA.
CC ProDom; PD000281; Gprotein_alpha; 1.
CC SMART; SM00275; G-alpha; 1.
CC GTP-binding; Transducer; Multigene family; ADP-ribosylation.
CC NP_BIND 46 53 GTP (BY SIMILARITY).
CC NP_BIND 205 209 GTP (BY SIMILARITY).
CC NP_BIND 274 277 GTP (BY SIMILARITY).
CC MOD_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX)
CC (BY SIMILARITY).
CC SEQUENCE 359 AA; 42026 MW; B4CD057E9FC7092A CRC64;
CC -----
Query Match 82.6%; Score 1510; DB 1; Length 359;
Best Local Similarity 80.6%; Pred. No. 4.4e-106;
Matches 283; Conservative 40; Mismatches 28; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRISAEIERHVRDCKDARRELKLLLTGSGSGKSTFIKQMRIHSGYS 64
Db 9 CCLDEVKESKRINAEIEKQLRDRDKRDARRELKLLLTGSGSGKSTFIKQMRIHSGYS 68

QY 65 DEDRKGFTKLVIYQNIFTAMQAMIRAMDTLRIQYMCQENKNAQIIREVEVDKVTALSRDQ 124
Db 69 EEDRKGFTKLVIYQNIFTAMQAVVRAMDTLKIRYKEQKANALLIREVDVEKVTTFEHQY 128

QY 125 VAAIKQLWLDPGIQECYDRRREYQLSDSAKYLLTDIERIAMPSSFVPTQQDVLVRVPTTG 184
Db 129 VNAIKTLWSDPGVQECYDRRREFQLSDSAKYLLTDVRIATVGYLPTQQDVLVRVPTTG 188

QY 185 IIEYFPDLENIIFRMVDVGGQSRERRKWIHCFSVTSIIFLVALSEYDQVLAECNENRM 244
Db 189 IIEYFPDLENIIFRMVDVGGQSRERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248

QY 245 EESKALFRTIITYPWFNLSSVILFLNKKDLLEEKIMYSHLISYFPPEYTGPKQDVKAARDF 304
Db 249 EESKALFRTIITYPWFQHSVILFLNKKDLLEDKILHSHLVVDYFPFEDGPPORDQAAREF 308

QY 305 ILKLYDQNPDKKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
Db 309 ILKMFVDLNPDSDKIISYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 15
GBQ_XENLA
ID GBQ_XENLA STANDARD; PRT; 353 AA.
AC P38410;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanine nucleotide-binding protein G(q), alpha subunit.
GN GNAQ.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBL_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Ladner K.J., Smith L.D.;
RL Submitted (XX-1992) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=94298961; PubMed=8026589;
RA Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;
RT "Neurexin B receptor, expressed in Xenopus laevis oocytes,
RT selectively couples to G alpha q and not G alpha 11.";
RL FEBS Lett. 348:89-92(1994).
RN [3]
RP ERRATUM.
RA Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;
RL FEBS Lett. 349:318-318(1994).
CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L05540; AAA49730.1; ALT_INIT.
CC EMBL; U10502; AAA52189.1; ALT_INIT.
CC HSSP; P10824; IBOF.
CC InterPro; IPR001019; Gprotein_alpha.
CC Pfam; PF00503; G-alpha; 1.
CC PRINTS; PR00318; GPROTEINA.
CC ProDom; PD000281; Gprotein_alpha; 1.
CC SMART; SM00275; G-alpha; 1.
CC GTP-binding; Transducer; Multigene family; ADP-ribosylation;
CC Palmitate; Lipoprotein.
CC LIPID 3 3 PALMITATE (BY SIMILARITY).
CC LIPID 4 4 PALMITATE (BY SIMILARITY).
CC NP_BIND 40 47 GTP (BY SIMILARITY).
CC NP_BIND 199 203 GTP (BY SIMILARITY).
CC NP_BIND 268 271 GTP (BY SIMILARITY).
CC MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX)
CC (BY SIMILARITY).
CC CONFLICT 10 10 E -> K (IN REF. 2).
CC CONFLICT 79 79 S -> T (IN REF. 2).
CC CONFLICT 158 158 L -> V (IN REF. 2).
CC CONFLICT 164 164 H -> Q (IN REF. 2).
CC SEQUENCE 353 AA; 41541 MW; 2310142412184553 CRC64;

Query Match 81.6%; Score 1492; DB 1; Length 353;
Best Local Similarity 80.1%; Pred. No. 9.5e-105;
Matches 281; Conservative 37; Mismatches 33; Indels 0; Gaps 0;

QY 5 CCLSAEEKESQRISAEIERHVRDCKDARRELKLLLTGSGSGKSTFIKQMRIHSGYS 64
Db 3 CCLSEEAEEARRINDEIERLRDRDKRDARRELKLLLTGSGSGKSTFIKQMRIHSGYS 62

QY 65 DEDRKGFTKLVIYQNIFTAMQAMIRAMDTLRIQYMCQENKNAQIIREVEVDKVTALSRDQ 124
Db 63 DEDRKGFTKLVIYQNIFSAMQAMIRAMETLKIPYKYEHNKGHALLVREVDVEKVASFENPY 122

QY 125 VAAIKQLWLDPGIQECYDRRREYQLSDSAKYLLTDIERIAMPSSFVPTQDVLVRVPTTG 184
Db 123 VDAIKYLWDPGIQECYDRRREYQLSDSTKYLLNDLDRIATHGYLPTQQDVLVRVPTTG 182
```


Qy	185	II	EY	P	D	L	E	N	I	I	F	R	M	V	D	V	G	G	R	S	E	R	R	K	W	I	H	C	F	E	S	V	T	S	I	I	F	L	V	A	L	S	E	Y	D	Q	V	L	A	E	C	D	N	E	N	R	M	244				
Db	183	II	EY	P	D	L	Q	S	V	I	F	R	M	V	D	V	G	G	R	S	E	R	R	K	W	I	H	C	F	E	N	V	T	S	I	M	F	L	V	A	L	S	E	Y	D	Q	V	L	V	E	S	D	N	E	N	R	M	242				
Qy	245	E	E	S	K	A	L	F	R	T	I	I	T	Y	P	W	F	L	N	S	S	V	I	L	F	L	N	K	K	D	L	L	E	E	K	I	M	Y	S	H	L	I	S	Y	F	P	E	Y	T	G	P	K	Q	D	V	K	A	A	R	D	F	304
Db	243	E	E	S	K	A	L	F	R	T	I	I	T	Y	P	W	F	Q	N	S	S	V	I	L	F	L	N	K	K	D	L	L	E	E	K	I	M	Y	S	H	L	V	D	Y	F	P	E	Y	D	G	P	Q	R	D	A	A	A	R	E	F	302	
Qy	305	I	L	K	L	Y	Q	D	Q	N	P	D	K	E	K	V	I	Y	S	H	F	T	C	A	T	D	T	E	N	I	R	F	V	F	A	A	V	K	D	T	I	L	Q	L	N	L	R	E	F	N	L	V	355									
Db	303	I	L	K	M	E	V	D	L	N	P	D	S	D	K	I	I	Y	S	H	F	T	C	A	T	D	T	E	N	I	R	F	V	F	A	A	V	K	D	T	I	L	Q	L	N	L	K	E	Y	N	L	V	353									

Search completed: February 3, 2003, 14:11:42
Job time : 14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: February 3, 2003, 14:04:16 ; Search time 35 Seconds
(without alignments)
2089.906 Million cell updates/sec

Title: US-09-492-028-2
Perfect score: 1828
Sequence: 1 MAGCCCLSAEEKESQRIAE.....VFAAVKDTILQLNLRFNLV 355

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp Vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archheap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1817	99.4	355	11 Q8R2X9	Q8r2x9 mus musculu
2	1524	83.4	359	4 Q9BZB9	Q9bzb9 homo sapien
3	1498	81.9	359	11 Q91X95	Q91x95 mus musculu
4	1435	78.5	353	5 P91955	P91955 limulus poi
5	1410	77.1	353	5 Q9U473	Q9u473 panulirus a
6	1408	77.0	355	5 Q17386	Q17386 caenorhabdi
7	1388	75.9	353	5 Q9NL92	Q9nl92 octopus vul
8	1385	75.8	353	5 Q8T6P8	Q8t6p8 mamestra br
9	1356	74.2	353	5 Q9NF20	Q9nfz0 calliphora
10	1352	74.0	360	5 Q8WSU8	Q8wsu8 halocynthia
11	1215	66.5	303	5 Q8T3G5	Q8t3g5 caenorhabdi
12	1123.5	61.5	305	5 Q9Y207	Q9y207 hydra magni
13	1065.5	58.3	355	5 Q9X2V4	Q9xzv4 geodia cydo
14	1028	56.2	279	5 Q9I7C8	Q9i7c8 drosophila
15	1011	55.3	374	6 Q9TU29	Q9tu29 oryctolagus
16	907.5	49.6	354	4 Q8TD72	Q8td72 homo sapien

17	892.5	48.8	354	13 Q90847	Q90847 gallus gall
18	891.5	48.8	354	4 Q9UGA4	Q9uga4 homo sapien
19	888.5	48.6	354	13 O13011	O13011 xenopus lae
20	884.5	48.4	354	13 Q90846	Q90846 gallus gall
21	883	48.3	355	13 Q9W6A4	Q9w6a4 squalus aca
22	882	48.2	355	4 Q96C71	Q96c71 homo sapien
23	879.5	48.1	354	4 Q8TAN5	Q8tan5 homo sapien
24	878.5	48.1	354	11 Q9DC51	Q9dc51 mus musculu
25	876	47.9	350	13 Q93565	Q93565 xenopus lae
26	875	47.9	305	5 Q9Y203	Q9y203 ephydatia f
27	872	47.7	353	3 Q96VA7	Q96va7 fusarium ox
28	872	47.7	353	3 Q96TI8	Q96ti8 gibberella
29	872	47.7	353	3 Q8TGA3	Q8tga3 leptosphaer
30	869	47.5	353	3 Q96VN7	Q96vn7 trichoderma
31	868.5	47.5	354	5 Q9NL93	Q9nl93 octopus vul
32	867.5	47.5	301	5 Q9Y202	Q9y202 ephydatia f
33	866	47.4	353	3 Q9HFA2	Q9hfa2 rosellinia
34	866	47.4	353	3 Q9C134	Q9cl34 tapesia yal
35	864.5	47.3	354	13 Q93564	Q93564 xenopus lae
36	863	47.2	371	13 Q8QGY8	Q8qgy8 fugu rubrip
37	861	47.1	357	5 Q9NL97	Q9nl97 halocynthia
38	855	46.8	469	5 Q8WPA0	Q8wpa0 halocynthia
39	852	46.6	353	3 Q9HEP9	Q9hep9 botrytis ci
40	850.5	46.5	354	5 Q8WP45	Q8wp45 halocynthia
41	848.5	46.4	354	3 Q9HFN1	Q9hfn1 blumeria gr
42	842.5	46.1	350	13 Q9DG28	Q9dgd28 gallus gall
43	840.5	46.0	354	5 Q9NL94	Q9nl94 octopus vul
44	840.5	46.0	354	5 Q8WSS1	Q8wssl ciona intes
45	838	45.8	354	3 Q9C115	Q9cl15 pisolithus

ALIGNMENTS

RESULT 1

Q8R2X9					
ID	Q8R2X9	FRELMINARY;	PRT;	355	AA.
AC	Q8R2X9;				
DT	01-JUN-2002 (TremBLrel. 21, Created)				
DT	01-JUN-2002 (TremBLrel. 21, Last sequence update)				
DT	01-JUN-2002 (TremBLrel. 21, Last annotation update)				
DE	Guanine nucleotide binding protein, alpha 14.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Strausberg R.;				
RL	Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL; BC027015; AAH27015.1; -				
SQ	SEQUENCE 355 AA; 41527 MW; D34B39ACD179AE82 CRC64;				

Query Match 99.4%; Score 1817; DB 11; Length 355;
Best Local Similarity 99.4%; Pred. No. 4e-136;
Matches 353; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1	MAGCCCLSAEEKESQRI	SAEIERHVRDKDARRELKLL	LLGTGESGKSTFIKQMRI	IHG 60
Db	1	MAGCCCLSAEEKESQRI	SAEIERLRDCKDARRELKLL	LLGTGESGKSTFIKQMRI	IHG 60
QY	61	SGYSEDRKGF	TKLVYQNIFTAMQAMIRAM	DTLRIQYCEQKENAIIRE	VEVDKVTAL 120
Db	61	SGYSEDRKGF	TKLVYQNIFTAMQAMIRAM	DTLRIQYCEQKENAIIRE	VEVDKVTAL 120
QY	121	SRDQVAAIKQLWLDP	GIQECYDRRREYQLSDS	AKYYLTDIRIAMP	SFVPTQQDVLVRV 180
Db	121	SRDQVAAIKQLWLDP	GIQECYDRRREYQLSDS	AKYYLTDIRIAMP	SFVPTQQDVLVRV 180
QY	181	PTTGIIIEY	PPFDLENIIFRMVDVG	QSRERRKWIHC	FESVTSIIFLVALUEYDQVLAECDN 240
Db	181	PTTGIIIEY	PPFDLENIIFRMVDVG	QSRERRKWIHC	FESVTSIIFLVALUEYDQVLAECDN 240

Db 9 CCPWCLSEDEKAAARVDQEITRLLLEHRRQVRGELKLLLLGTGESGKSTFIKQMRIIHGA 68

QY 62 GYSEDRKGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALS 121

Db 69 GYSEEDRKGRPLVFNQIFLSVQAIIEAMDRLQIPYSRPESKLHASLVMSQDPYKVNTFE 128

QY 122 RDQVAAIKQLWLDPGIQECYDRRREYQLSDSAKYIYLTDIERIA MPSFVPTQQDVLRVRVP 181

Db 129 TRYALAVQSLWRDAGVRACYERRREFHLLDSAVYILSHLERIAEEGYVPTAQDVLRSRMP 188

QY 182 TTGIIEYPFDLENIIFRMVDVGGQSRERRKWIHCFESVTSIIFLVALSEYDQVLAECDNE 241

Db 189 TTGINEYCFSVQKTNLRIVDVGGQKSERRRKWIHCFEDVTALIFLASLSEYDQCLEENGQE 248

QY 242 NRMEESKALFRTIITYPWFLNSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAA 301

Db 249 NRMQESLALFGTVLALPWFRATSVILFLNKTDLIEDKVRTSHLATYFPGFRGPPQDPEAA 308

QY 302 RDFILKLY-----QDQNPD-----KEKVIYSHFTCATDTENIRFVFAAVKDTILQLNL 349

Db 309 KRFILELYTRYAGAAAGPDGASKGPRSRRLFSHYTCATDTQNIKVKFDVRDSVLARYL 368

QY 350 REFNLV 355

Db 369 DEINLL 374

Search completed: February 3, 2003, 14:11:22

Job time : 37 secs